Titre: Title:	A Framework for the Performance Analysis and Simulation of RF-Mesh Advanced Metering Infrastructures for Smart Grid Applications
Auteur: Author:	Filippo Malandra
Date:	2016
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Malandra, F. (2016). A Framework for the Performance Analysis and Simulation of RF-Mesh Advanced Metering Infrastructures for Smart Grid Applications [Ph.D. thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/2422/

UNIVERSITÉ DE MONTRÉAL

A FRAMEWORK FOR THE PERFORMANCE ANALYSIS AND SIMULATION OF RF-MESH ADVANCED METERING INFRASTRUCTURES FOR SMART GRID APPLICATIONS

FILIPPO MALANDRA
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(GÉNIE ÉLECTRIQUE)
DECEMBER 2016

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

A FRAMEWORK FOR THE PERFORMANCE ANALYSIS AND SIMULATION OF RF-MESH ADVANCED METERING INFRASTRUCTURES FOR SMART GRID APPLICATIONS

présentée par: MALANDRA Filippo
en vue de l'obtention du diplôme de: Philosophiæ Doctor
a été dûment acceptée par le jury d'examen constitué de:

M. GIRARD André, Ph. D., président
Mme SANSÒ Brunilde, Ph. D., membre et directrice de recherche
M. SIROIS Frédéric, Ph. D., membre
Mme KASSOUF Marthe, Ph. D., membre externe

DEDICATION

To my father,
may he be proud of me
as much as I am proud of him.
Ciao Papì...

ACKNOWLEDGEMENTS

First of all, I would like to thank my research advisor Professor Sansò, who wanted me to come back from Italy after my previous experience in 2010 : it's mainly thanks to her that I moved to Montreal and I do not regret at all my decision. Then, I would like to thank the SmartDESC project and all the involved students and professors from whom I learned a lot : a particular mention to Michel Sauvé for his passion, his inspiring life stories, and for the motivation he was able to transfer to me in a particularly delicate moment of my life, when I was in Italy and I thought I would never be able to come back and pursue my Ph.D. program. Then, I would like to thank the members of my jury for accepting to review my thesis and to play a very important role in this last stop of my Ph.D. journey.

I would also like to thank all the wonderful travelmates, with whom I shared this intense 4-year trip. My first thought is for my father, who unfortunately witnessed only the beginning of it : he is my life-mentor and what I learned from him will certainly accompany me for the rest of my life, along with all the wonderful memories permanently stored in my heart. I want to thank my whole family for the strength they are able to give me, for their unconditional love and their immense trust on my capabilities : my mum Luciana, who taught me the value of sacrifice for the people you love (as well as the art of cooking); my sister Mariassunta, who has always been a model for me with her energy and joy; my brother Andrea, who taught me the importance of being rational and balanced in every situation, even the most difficult ones; my cousin Nicola, who taught me the importance of the smile against life adversities and bad people as well as introducing me to the wonderful world of A.C. Milan; my sister-in-law Fabiana and my brother-in-law Francesco who built two wonderful and happy families with my brother and sister; and all other uncles, aunts, and cousins, who are active part of my big family. I would like to especially thank my marvelous nieces Lulu and Cici, and my awesome nephew Lollo (not to mention the forthcoming Ceciotto) : they are the light in my eyes. My love for them did not fade of a single dB in the thousands of kilometers that currently separate us. The same applies to my lifelong friends Paolo, Armando and Andrea : may our friendship never end.

A special thank goes to my football team *Saskatchewan* and to all the amazing teammates : Luca, Davide, Charles, Alessandro, Matteo, Alejandro, Alexandre, Ken and Agustin just to cite some of them, with whom I played, won, lost, and (most importantly) had a lot of fun. Moreover, I got in contact with a wonderful community of young Italian "emigrants" like me (Luca, Matteo, Giovanni, Eleonora, Margherita, Franco, Lidia, Alessandro, Alice,

Davide, Valentina, Carmelo, Francesco, and many many others) : with them I shared the wanderlust, the homesickness, the passion for football and for good food, some amazing trips to explore the wonderful country that hosts us, and many cups of coffee (good short Italian coffee, of course). Anyways, I did not only get along with Italian folks, but I also made new friends from all over the world : from Argentina to Netherland, from France to Spain, from Swiss to Greece, from Albania to Colombia, from Brazil to Egypt, from Tunisia to Morocco, from Syria to China, from Turkey to Iran, from Lebanon to Venezuela, from Mexico to Canada, and many others. I am infinitely grateful to them all for giving me the stability and happiness I needed to cope with the absence of family and the nostalgia.

A special thank goes to Ecole Polytechnique de Montreal and GERAD, two amazing entities who provided me with everything I could need in this trip. At Polytechnique, I would like to especially thank Nathalie, for helping me to go through the jungle of the bureaucracy I crossed on my way, and Rachel, to whom I am very grateful for the precious help and for the kindness and courtesy she showed me in every occasion. I would also like to thank all the personnel at GERAD, and in particular Marie, Edoh and Pierre : they are very competent and have always been ready to help me in every kind of problems I encountered during my research. The GERAD is a very positive and powerful environment to study and work, thanks to the many students from many countries around the world : I am grateful for being a part of this multicultural research community.

At last but not least comes Loli, my girlfriend, life partner and best friend. I cannot express with words my gratitude for all she did for me (I should probably write another thesis about that...) but I want to thank her for being there with me in this experience far from our families, for deciding to leave everything behind and follow me in this adventure. I want to close with just one italian word that comes out of my heart, GRAZIE!

RÉSUMÉ

L'Infrastructure de Mesurage Avancée (IMA), conçue à l'origine pour lire à distance des compteurs intelligents, est actuellement considérée comme une composante essentielle dans le domaine des Smart Grid. Le but principal des IMAs est de connecter le grand nombre de compteurs intelligents installés chez les clients au le centre de contrôle de données de l'entreprise d'électricité et viceversa. Cette communication bidirectionnelle est une caractéristique recherchée par un grand nombre d'applications, qui visent à utiliser ces infrastructures comme support à la transmission de leurs données dans le réseau électrique, comme par exemple la gestion de la charge et la demande-réponse. Un grand nombre de technologies et de protocoles de communication sont actuellement utilisés dans les IMAs : parmi les solutions disponibles, le RF-Mesh est une des plus populaires, surtout grâce au bas coût pour l'installation et les équipements. Toutefois, le débit nominal des communications RF-Mesh est très bas, de l'ordre des dizaines de kbps, et la littérature qui traite leur performance est très limitée. Ceci pourrait en limiter l'utilisation pour des applications autres que la lecture à distance des compteurs intelligents.

Ce travail de thèse vise à développer un système de modèles et outils pour évaluer la performance des réseaux RF-Mesh et encourager leur utilisation pour un grand nombre d'applications dans le domaine des Smart Grid. Le système d'évaluation de performance proposé est constitué (i) de modèles analytiques, pour calculer la probabilité de collision entre les paquets transmis, (ii) d'un simulateur de réseau, pour recréer le fonctionnement des réseaux RF-Mesh dans un environnement virtuel, (iii) d'un générateur de topologie, pour créer des cas réalistes en se basant sur des données géographiques et (iv) des méthodes pour l'analyse de la performance.

Trois différents modèles analytiques ont été implémentés. Dans les deux premiers, une nouvelle formule analytique a été utilisée pour calculer la probabilité de collision entre paquets. La probabilité de collision est ensuite utilisée pour estimer le délai moyen de/vers chaque compteur intelligent dans l'IMA analysée. Par la suite, des indices de performance, basés sur le délai moyen, sont utilisés pour faire des analyses de performance : études de faisabilité pour les applications de Smart Grid, l'identification de nœuds critiques et d'éventuels goulots d'étranglement. Dans le troisième modèle analytique, la théorie de Markov-Modulated System est utilisée pour prendre en considération d'importants détails d'implémentation, comme la probabilité de retransmission et la taille des mémoires tampons des nœuds, qui n'avaient pas été inclus dans la modélisations précédente. Ce dernier modèle fournit aussi un mode

de calcul plus précis de la probabilité de collision en raison de la plus grande complexité de modélisation, par rapport aux deux modèles analytiques précédents. Un trafic distribué selon la loi de Poisson est utilisé pour représenter le trafic généré par plusieurs types d'applications Smart Grid.

Le cadre dévaluation de la performance proposé inclut aussi un simulateur de réseau, qui a été implémenté à partir de zéro en utilisant Java et Python. L'outil de simulation permet de modéliser des parties du système, comme le routage dynamique ou des différentes distributions de trafic, qui sont difficiles à reproduire par un modèle mathématique. Des résultats numériques sont ainsi présentés pour montrer les différents types d'analyse qui peuvent être réalisées avec les outils proposés. Les résultats obtenus par le modèle analytique et par le simulateur de réseau ont été comparés pour montrer la cohérence entre les deux approches poursuivies dans cette thèse. En plus, le système pour l'analyse de performance, qui est objet de cette thèse, a été utilisé dans le cadre du projet Smart Distributed Storage Energy Controller (SmartDESC), un projet multidisciplinaire de gestion de la charge électrique. Les détails théoriques de l'intégration de l'outil de simulation ont été présentés, ainsi que de résultats numériques, pour montrer une des possibles applications pratiques du contenu de cette thèse.

ABSTRACT

Advanced Metering Infrastructure (AMI), originally conceived to replace the old Automated Meter Reading (AMR) infrastructures, have now become a key element in the Smart Grid context and might be used for applications other than remote meter reading. The main driver to their widespread installation is that they provide power utilities with a bidirectional connectivity with the smart meters. A wide variety of communication networks are currently proposed to support the implementation of AMIs, and, among them, the RF-Mesh technology seems to be very popular. The main reasons for its adoption are the proprietary infrastructure and the modest cost for the installation and the equipment. However, RF-Mesh systems are characterized by poor achievable data-rates in the order of 10 kbps, and their performance is not well studied in the literature. The lack of tools and methods for the performance evaluation might be a roadblock to their widespread adoption. This thesis aims at filling this gap and increase the knowledge of large-scale RF-Mesh systems to foster their use for a wide variety of applications.

We propose a comprehensive framework for the performance evaluation of large-scale AMIs adopting the RF-Mesh technology. The framework includes (i) a geo-based topology generator that uses geographic data to produce realistic AMI cases, (ii) analytic models for the computation of packet collision probability and delay, (iii) a network simulator to recreate the behavior of large-scale RF-Mesh systems, and (iv) methods to evaluate the performance.

Three different analytic models are included in the framework. The first two provide a novel analytic formulation of the packet collision probability in a mesh network with time-slotted ALOHA and the Frequency Hopping Spread Spectrum (FHSS) protocol : the collision probability is then used to estimate the average delay in the network, and to define and evaluate performance indexes (e.g., critical nodes and survival function). In the third model, a complex Markov-Modulated System (MMS) is used to take into consideration important implementation details, such as the retransmission probability and the buffer size, that were not considered in the two previous models. This model also provides a more accurate computation of the packet collision probability. A Poisson distribution is used to represent the traffic coming from potential Smart Grid applications. The framework also includes an RF-Mesh network simulator, written in Java and Python. The tool provides additional enhanced features with respect to the analytic models, such as a dynamic routing protocol or different traffic distributions.

Numerical results are provided to show the type of performance evaluations that is possible using the proposed framework. A comparison between simulation and analytic results is also proposed, showing consistency between the two models. Finally, the performance evaluation framework was applied in the context of Smart Distributed Storage Energy Controller (SmartDESC), a multi disciplinary load-management project. Details about the theoretical aspect of the integration of the implemented network simulator and the SmartDESC simulator are proposed along with numerical results.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AMI Advanced Metering Infrastructure

AMR Automated Meter Reading

AODV Ad-hoc On-demand Distance Vector

DAG Directed Acyclic Graph

DCU Data Collection Unit

DoE Department of Energy

DSR Dynamic Source Routing

FERC Federal Energy Regulatory Commission

EV Electric Vehicle

EWH Electric Water-Heater

FHSS Frequency Hopping Spread Spectrum

GIS Geographic Information System

GUI Graphical User Interface

HAN Home Area Network

ICT Information, Communication, and Technology

IMA Infrastructure de Mesurage Avancée

IoT Internet of Things

ISM Industrial, Scientific, and Medical

LPWAN Low Power Wide Area Network

MAC Media Access Control

MDMS Metering Data Management System

MMS Markov-Modulated System

NAN Neighborhood Area Network

NIST National Institute of Standards and Technologies

ORPL Opportunistic RPL

PLC Power Line Communications

RFI Request For Information

RMSE Root Mean Square Error

RPL Routing Protocol for Low-power and lossy networks

SmartDESC Smart Distributed Storage Energy Controller

V2G Vehicle-to-grid

WAN Wide Area Network

WMN Wireless Mesh Network

WSN Wireless Sensor Network

LIST OF APPENDICES

CHAPTER 1 INTRODUCTION

1.1 Context and motivations

The traditional power grid, with its old and centralized architecture, is incapable of satisfying the power needs of the modern society, with the global population constantly increasing and a large number of electronic devices available at an affordable price. As a consequence, the last decades have witnessed the rise of the *Smart Grid* concept, that renews and restructures the energy sector : the main idea is to make the existing power grid smarter and able to keep pace with the technological breakthroughs in other sector.

The introduction of smart meters paved the road for the implementation of smart grids on a large scale. Smart meters were conceived to perform the remote reading of the power consumption at the customer premises. They were only intended to replace the old mechanic power meters, which were not able to communicate and required periodic human reading. The first kind of communication infrastructure among smart meters was the so-called Automated Meter Reading (AMR) network, which was put in place by power utilities to remotely read the reader without sending an operator at all the premises. However, the unidirectional flow of information provided by AMR communication networks soon became insufficient to cope with the smart grid paradigm, for which the two-way communication between the power utility and the endpoints of the grid is fundamental.

AMIs were proposed to go beyond the limits of AMR systems and provide bidirectional connectivity between the smart meters and the power utility Metering Data Management System (MDMS). In Figure 1.1, the total number of installed meters in U.S. is shown for both AMR and AMI. The graph shows how the installation of AMI meters in U.S. has progressively replaced the installation of new AMR meters. The two-way AMI connectivity allows the power utility to send feedback to the customers (e.g., to change the price according to the availability of energy), to control devices at distance (e.g., water-heaters or space heaters), to perform load-management, to enhance the monitoring and awareness of all the elements of the grid (e.g., substations, smart meters). The peculiarities of AMIs are the high number of nodes, the data confidentiality, the large geographic extension of the topology (e.g., in rural areas), and the high concentration of nodes (e.g., in urban areas). Wireless solutions prevail over wired networks because of their flexibility, reduced equipment cost, and ease of installation. A wide variety of wireless technologies and protocols has been proposed so far to support AMI : the most widespread are ZigBee, WiFi, 802.15.4, 6LoWPAN, cellular (2G, 3G, and 4G), satellite, and RF-Mesh.

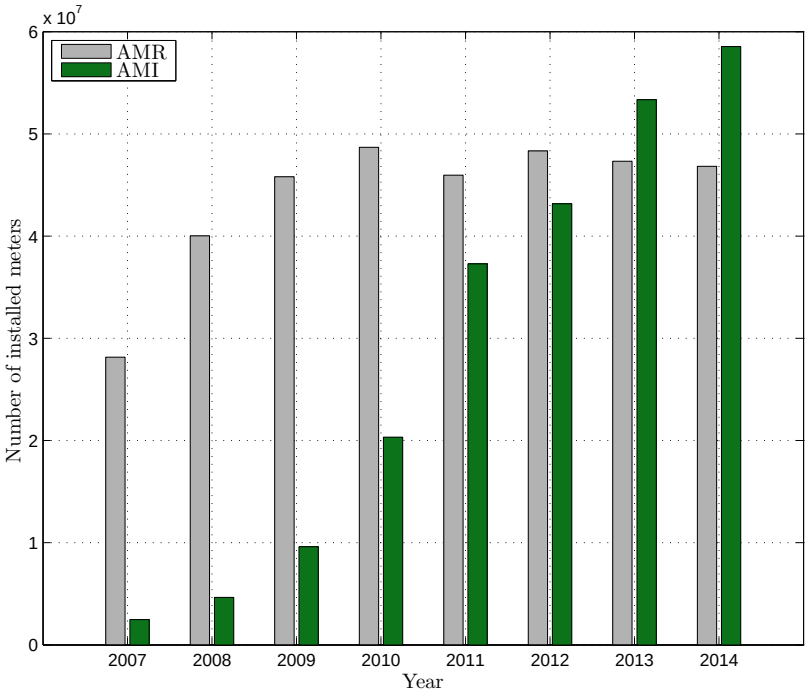


Figure 1.1 Trend of the total number of installed meters for AMR and AMI in U.S. from 2007 to 2014. Data extracted from : `http://www.eia.gov/electricity/annual/html/epa_10_10.html`.

RF-Mesh, despite the low achievable data-rates, is becoming particularly popular for AMI and is currently being considered by a large number of smart meter manufacturers as a built-in solution in their devices. Suffice it to say that Aclara, Itron, and Landys&Gyr — smart meters manufacturers that own 76.5% of the market share in North America, according to Navigant Research (2016) — install RF-Mesh radio in their systems[1]. RF-Mesh AMI are cost-effective and use a proprietary infrastructure, which drastically reduces the dependence of power utilities on telecommunication providers : this is why RF-Mesh systems are being currently considered as a valuable AMI solution.

Despite their popularity, the performance of RF-Mesh has not been thoroughly studied in the current literature. One of the main causes is the lack of data : RF-Mesh systems are usually installed as *black-boxes* and many implementation details are covered by strong confidentiality agreements, which undermine their public dissemination. The lack of a thorough RF-Mesh performance literature considerably restrain their use for purposes other than remote reading : the introduction of new applications in a communication system needs a preliminary

[1]For more information, `http://www.greentechmedia.com/articles/read/who-are-the-top-ten-vendors-in-smart-grid/all`

performance evaluation of the system, to see if it is capable of meeting thee communication requirements of the applications.

1.2 General Objectives and Original Contribution

As highlighted in Section 1.1, RF-mesh systems are more and more widespread in the context of AMI, and an increasing number of applications could benefit from its use. It is therefore the main objective of this thesis to provide a *framework for the performance analysis* of RF-Mesh systems. For this, we developed insight on the following issues.

- **Large-scale instances**

  RF-Mesh AMI are composed of several thousands of smart meters. The large number of nodes considerably affect the analysis because it increases the computational burden associated to both simulations and analytic models. It is important to propose methods and models that are computationally efficient and able to work with thousands of nodes.

- **Wireless environment prone to collision**

  The free and unlicensed Industrial, Scientific, and Medical (ISM) bandwidth entails high chances of packet collisions. The Colliding packets need to be retransmitted, increasing the delay and degrading the performance. A correct model for the collision probability is key to an accurate representation of the RF-Mesh systems. The collision probability calculation needs to consider the presence of the FHSS protocol.

- **Geographical distribution of nodes**

  A RF-Mesh topology includes smart meters, routers, and collectors, that are far from being randomly distributed in space. In particular, smart meters are installed at residential and commercial premises, which follow peculiar spatial distributions. The use of geographic data to build up the topology permits more realistic instances and produces a more accurate analysis.

- **Black-box nature**

  RF-Mesh implementation details are usually covered by strict confidentiality agreements and not available to the public. The difficulty in retrieving valuable data jeopardizes the development of performance studies and needs to be addressed by means of what-ifs analyses and reasonable assumptions.

- **Low achievable throughput**

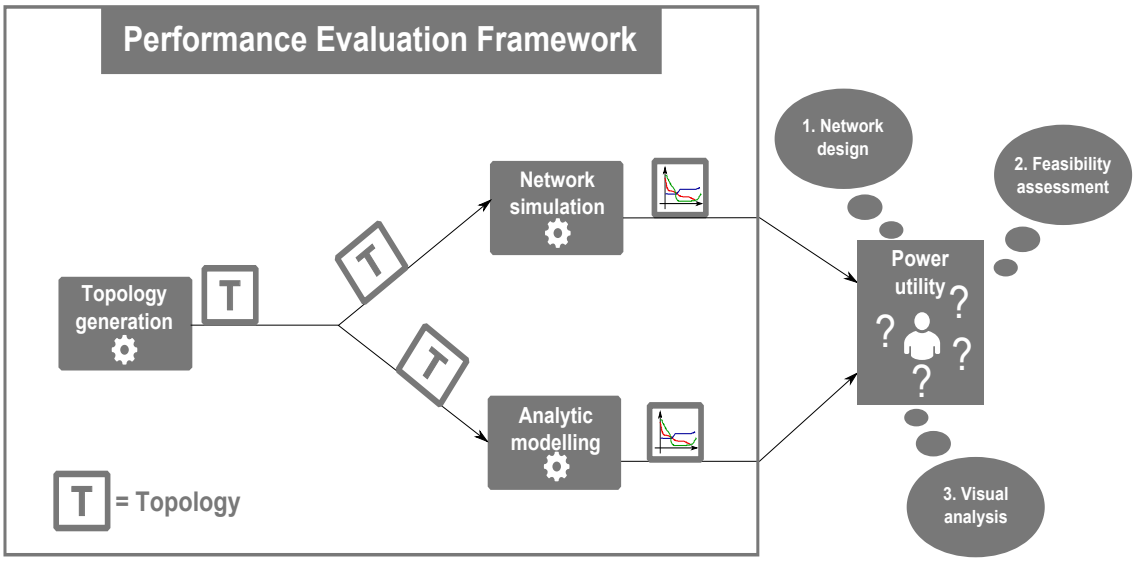


Figure 1.2 Simplified architecture of the proposed performance evaluation framework.

The nominal capacity of wireless links is in the order of tenths of kbps, which does not seem to keep up with current and more advanced communication system. The 4G, for instance, can ideally achieve 100 Mbps, roughly ten thousands time faster than RF-Mesh. The low achievable datarates need to be considered in the analysis.

- **MAC layer**

  The access to the communication medium, which is managed according to the slotted ALOHA protocol, plays a key role in the performance evaluation because it largely affects the packet collision probability calculation.

- **Network layer**

  The network layer, and in particular the choice of the routing protocol, are very important in the definition of the network performance. Even though details about the routing protocols are not always available, it is fundamental to include the routing in the performance evaluation framework because it has an impact on the packet delay.

- **Lack of data** Given the scarcity of performance studies in the literature, validation must be carried out through simulations and analytical studies.

In order to cope with these requirements, a performance evaluation framework was proposed to analyze large-scale RF-Mesh systems. The structure of the proposed framework is illus-

trated in Figure 1.2 showing the following possible usages, from a power utility's perspective :

1. the design of a new RF-Mesh system;
2. the feasibility assessment of potential applications in a given infrastructure;
3. the visual identification of potential bottlenecks.

The following major original contributions have been produced during the doctoral program :

1.2.1 Performance analysis tools

Three different mathematical models were proposed to find analytic equations to calculate variables representing the state of the system (e.g., collision probability, delay) according to relevant input parameters (e.g., topology, traffic, routing), as shown in the prototype architecture in Figure 1.3.

The first model, described in detail in Chapter 4, proposes an expression for the collision probability and the delay, with a basic interference model not considering the FHSS protocol. The second model, discussed in Chapter 5, introduces and analyzes the impact of FHSS on the collision probability and the delay. In the third model, described in Chapter 7, a more complex Markov-Modulated System (MMS) was used in order to include in the analysis important implementation details, such as the buffer size and the retransmission probability, not included in previous models.

1.2.2 Network simulator

A RF-Mesh network simulator, described in Chapter 6, was developed from scratch using Java and Python to obtain numerical results to be compared to the analytic results. The simulator proved to be computationally efficient allowing the evaluation of several thousand node instances in a reasonable time (few hours at maximum). The simulator yields the delay and other performance indexes, enabling a wide variety of performance analyses, as reported in Section 6.4. Moreover, the user can choose between two types of routing, shortest path and layer-based. Two different types of traffic are implemented : (i) Poisson-distributed (as in the analytic), and (ii) deterministic (e.g., scheduled broadcast transmissions at the collector side). Simulation results were compared to the analytic results, as highlighted in Section 7.5.5, showing consistency between the two different approaches.

1.2.3 Geo-based topology generator

The tool, described in Chapter 8, produces realistic topologies that are used throughout the work. Publicly available datasets were employed to retrieve the position of smart meters, whilst the position of routers and collectors was chosen in a map, according to the data published in Hydro-Québec (2012).

1.2.4 Dedicated simulation module for the SmartDESC project

A modified version of the network simulator was developed to be used in the SmartDESC project, a load-management project whose main purpose was to optimally control several millions of electric water heaters installed in Québec. The telecommunication module, described in Chapter 9, was needed to emulate the behavior of an AMI, to study its performance, and to verify whether it was able to meet the communication requirements (e.g., the delay) of the SmartDESC project.

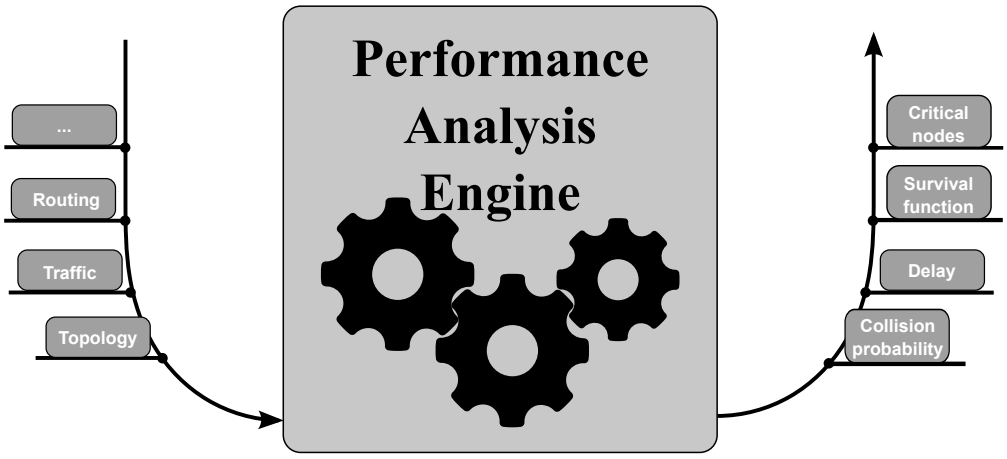


Figure 1.3 Workflow of the performance analysis engine.

CHAPTER 2 ORGANIZATION OF THE RESEARCH WORK

This thesis is structured *by articles* and is composed of 4 papers :

- Published in peer-reviewed conference proceedings :

  1. F. Malandra, B. Sansò. *Analytical performance analysis of a large-scale RF-mesh smart meter communication system.* Innovative Smart Grid Technologies Conference (ISGT), 2015 IEEE Power & Energy Society, 1-5, Washington, 2015. Presented in Chapter 4.
  2. F. Malandra, B. Sansò. *PeRF-Mesh : A performance analysis tool for large scale RF-mesh-based smart meter networks with FHSS.* 2015 IEEE International Conference on Smart Grid Communications (SmartGridComm), 792–797, Miami, 2015. Presented in Chapter 5.

- Submitted to peer-reviewed journals :

  3. F. Malandra et B. Sansò, (2016), *A Simulation Framework for Network Performance Evaluation of Large-Scale RF-mesh AMIs.* Simulation Modelling Practice and Theory, pp. 26. Presented in Chapter 6.
  4. F. Malandra et B. Sansò, (2016), *A Markov-modulated End-to-end Delay Analysis of Large-scale RF-Mesh Networks with Time-slotted ALOHA and FHSS.* IEEE transactions on wireless communications, pp. 30. Presented in Chapter 7.

The main contribution of this thesis are :

- three mathematical models, for performance evaluation and feasibility assessment
- a network simulator, implemented from scratch, for performance evaluation and result validation
- a topology generator, that produces large-scale instances based on Geographic Information System (GIS) data
- a dedicated simulation module for the telecommunications used in SmartDESC, a load-management project partially funded by Natural Resources Canada

The thesis is essentially divided in five main parts :

- Introduction and literature review
- Modelling theory (analysis and simulation)
- Practical implementations and case study
- Conclusion and general discussion
- Annexes

The first part is composed of Chapters 1 and 3 : Chapter 1 includes the introduction to the subject, the context, the motivation, the research objectives and the main achievements of this thesis; Chapter 3 contains a critical review of the literature.

The second part is composed of the four papers, each included in a dedicated chapter. The three mathematical models are presented in Chapter 4, 5, and 7. In the first model, a large-scale RF-Mesh system is analyzed. An analytic formulation is proposed to calculate the probability of collision. Based on that, the multi-hop delay is computed and some performance indexes are defined and used in the performance analysis. The system under study is characterized by a shortest path routing, a Poisson-distributed traffic in the uplink and in the downlink, and a time-slotted ALOHA MAC. However, the FHSS protocol was not included in the modelling. As a consequence, the results proved the necessity of adopting a protocol to mitigate interference, such as FHSS. In the second model, presented in Chapter 5, the analyzed system is akin to that considered in Chapter 4, but the FHSS protocol was included in the collision probability formulation. The results showed the high impact on the performance of the FHSS protocol. Even though the first two analytic models allow valuable performance analyses of RF-Mesh systems, several implementation details were not included in the modelling, such as the retransmission probability and the buffer size. An enhanced analytic model, that aims at filling this gap, is proposed in Chapter 7. A MMS modelling was used to represent the state of each node at each time interval. This permits to have more accuracy in the collision probability calculation. For instance, the state of the system not only accounts for the transmission probability but also considers the destination of the transmission : this considerably affects the collision probability and produces more accurate results, as confirmed in the validation analysis proposed in Section 7.5.5. Finally, a RF-Mesh network simulator was conceived and implemented from scratch, using Java and Python. The simulator, presented in Chapter 6, proposes a completely different approach to the RF-Mesh performance evaluation with respect to the previously mentioned analytic models. In particular, the packet delay is simply calculated as the difference between the arrival time and the generation time of each packets. The network simulator permits to

choose between a shortest path routing, as in the analytic modelling, and the more dynamic *layer-based routing*.

The third part includes the implemented topology generator and a case study. The topology generator is employed to produce the instances in all the aforementioned papers and is described with detail in Chapter 8. In Chapter 9, a smart grid case study is presented to show the capabilities of the proposed RF-Mesh performance evaluation framework.

The fourth part of the thesis is composed of Chapter 10, which contains the concluding remarks and presents the synthesis of the work, and of Chapter 11, which includes the a discussion on the limitation of the proposed solution and some recommendations about future work.

The last part includes two appendices : Appendix A provides some additional detail on the Poisson-distributed traffic characterization, which is used throughout this thesis; Appendix provides additional mathematical proofs concerning the model presented in Chapter 7.

CHAPTER 3 LITERATURE REVIEW

3.1 Smart Grid and telecommunications

The Information, Communication, and Technology (ICT) sector can be considered a fundamental resource to facilitate the advent of the Smart Grid. The role of ICT in this context is investigated by a large branch of the literature. Recent surveys on this topic can be found in Erol-Kantarci and Mouftah (2015), Khan et al. (2016), El-Hawary (2014), and Farhangi (2014). A complete overview of the recent research developments can be found in Colak et al. (2015), Tuballa and Abundo (2016), and Bayindir et al. (2016). Other interesting surveys, useful to understand the leading objectives envisioned for the next years, are Hassan and Radman (2010), El-Hawary (2014), Farhangi (2014), and Hossain et al. (2016).

A Smart Grid requires the coexistence and integration of several networks and entities, making a standardization process fundamental. The National Institute of Standards and Technologies (NIST) was one of the first to put effort into this theme, presenting a framework to coordinate the wide range of involved actors in the domain of Smart Grid NIST (2010). Lo and Ansari (2012) carried out a combined analysis of the power grid and the communication infrastructure, discussing the interconnections between the two systems; in particular they reviewed the main Smart Grid applications, highlighting their communication requirements. In US DoE (2010a), a very interesting report published by the U.S. Department of Energy (DoE) in cooperation with the Federal Energy Regulatory Commission (FERC) and the NIST, six prime Smart Grid functionalities were identified together with their communication needs : (i) AMI, (ii) demand-response, (iii) electric vehicles, (iv) wide-area situational awareness, (v) distributed energy resources and storage, and (vi) distribution grid management. The rest of this chapter focuses on AMIs since it is the object of this thesis.

3.2 AMI architectures and technologies

AMIs are widely employed in Smart Grids, because they provide bidirectional communication between the power utility and the end points of the grid (i.e., the smart meters). An overview of current AMI deployments around the world can be found in Kabalci (2016), Lopez et al. (2015), Abdulla (2015), Chren et al. (2016), and Chakraborty and Shaniia (2016). Details on the recent initiatives and the ongoing projects involving AMIs can be found in Renofio et al. (2016), Garcia-Hernandez (2015), and Foudeh and Mokhtar (2015).

The communication requirements of AMI systems are strictly dependent on the type of application that is sought and on its communication requirements. In 2010, the US DoE also published US DoE (2010b), a Request For Information (RFI) to gather the communication requirements identified by several actors involved in the Smart Grid and to ease the standardization process : power utilities, telecommunication providers, wireless device manufacturers and customers representatives were asked to express their opinion and recommendations. In US DoE (2010a), the replies to the RFI are reported : the interviewed companies published their communication needs (e.g., delay tolerance and service reliability) associated to different Smart Grid applications (e.g., home energy management and demand-response). Gungor et al. (2013) provided additional data about the communication requirements for different applications. According to the authors, many applications (demand-response, load-control, dynamic pricing, billing, Vehicle-to-grid (V2G) and Electric Vehicles (EVs) charging) seemed to be delay-tolerant and potentially suitable for RF-mesh AMIs.

The most common types of AMI have a layered architecture, generally composed of a Home Area Network (HAN), which includes the short-range communications within a house, a Neighborhood Area Network (NAN), which relays data from several houses to a collector, and a Wide Area Network (WAN), an IP-backbone connecting all the data collectors to the MDMS, in charge of analyzing the smart meter data. Several technologies and protocols are adopted in AMIs and can be classified according to the layer of the architecture where they are used : more information about HAN, NAN, and WAN can be found in Noorwali et al. (2015), Ramirez et al. (2015), and Ali et al. (2016) respectively.

Gungor et al. (2011) proposed a critical review of the existing communication technologies used in the Smart Grid, describing their main features, benefits, limitations and range of application. Multiple papers study AMIs adopting wireless technologies that are considered standards, such as WiFi (e.g., in Tonyali et al. (2016), Sousa et al. (2015), Hu et al. (2015), Ramirez et al. (2015), and Michaloliakos et al. (2016)), cellular networks (e.g., GSM in Das and Saikia (2015), 3G in Athanasios and Cottis (2016), and LTE in Carlesso et al. (2015) and Yaacoub and Kadri (2015)), ZigBee (e.g., in Somkaew et al. (2014), Parvez et al. (2016), Chang et al. (2015), Peng and Huang (2016), and Chi et al. (2016)), RF-Mesh (e.g., in Lichtensteiger et al. (2010)), Z-wave (e.g., in Alliance (2011), Zareei et al. (2011), and Amaro et al. (2011)), Bluetooth (e.g., in Bluetooth (2010) and Wang et al. (2013)), and Low Power Wide Area Networks (LPWANs) (e.g., LoRa in Vangelista et al. (2015)).

3.3 Performance studies for AMIs

An increasing percentage of research proposes performance studies for wireless AMIs : recent surveys on this topic can be found in Mohassel et al. (2014), Sharma and Saini (2015), and Namboodiri et al. (2012). The existing body of literature proposes performance studies focusing on several specific aspects of AMIs, such as security, resilience, physical and network layer protocols. The different topics are treated in the rest of this section.

3.3.1 Security

The theme of security is central in many studies (Anzalchi and Sarwat (2015), Bhatia and Bodade (2014), Somkaew et al. (2014), Soykan et al. (2015), and Tonyali et al. (2016)) because AMIs usually deal with confidential data, that need to be protected. Accordingly, smart meter manufacturers have to adopt severe hardware counter-measures to unexpected intrusion in their devices. AMIs are prone to several types of attacks. One of the most common is the so-called *eavesdropping*, in which the intruder intercepts some traffic, decrypts it, and gains access to confidential data : as highlighted in Grochocki et al. (2012), this type of attack may result in the theft of sensitive customer information and can be blocked by adopting more robust encryption techniques. Another security threat is brought by the so-called *distributed denial of service* attack, analyzed in Jin et al. (2011) : this type of attack is usually directed to Data Collection Units (DCUs) and consists in tampering the regular operation of the smart meters by installing malware, which is used to flood the DCU with a large number of malicious packets. A good countermeasure to this attack consists in using stronger authentication mechanisms, that prevent unauthorized users from controlling the smart meters. A third possible AMI vulnerability might be the fraudulent use of the connect/disconnect feature : the intruder can take control of a DCU, and unexpectedly disconnect some smart meters from the grid, without being detected by the power utility. Two of the most widespread solutions to the aforementioned issues are advanced data encryption techniques, as discussed in Somkaew et al. (2014); Soykan et al. (2015); Tonyali et al. (2016), and strong authentication mechanisms, as presented in Lee et al. (2014) and Thomas et al. (2012).

3.3.2 Resilience

The resilience of AMI networks can be defined as their capacity to resist to unexpected node or link failures. Some authors decided to focus on the topology design in order to assure network resilience. Hartmann et al. (2014) proposes a generator of random but realistic

topologies, according to some resilience parameters, such as the maximum number of smart meters per concentrator or the maximum number of hops between each smart meter and the concentrators. Another methodology for the topology design was proposed by Renofio et al. (2016), and consists in choosing the location of gateways in a given topology in order to assure a minimum number of independent paths from the gateway to each smart meter. An interesting study of the resilience of AMIs using WiFi is Aruzuaga et al. (2010), where the authors propose the employment of a mixed single-hop multi-hop architecture to enhance the network resilience. In Galli et al. (2011), graph theory is exploited to measure the network resilience and the ability to resist to malicious attacks.

3.3.3 Physical layer

AMIs are characterized by a large number of nodes sharing the same physical resources (e.g., bandwidth, number of channels). Moreover, the use of additional resources is usually very expensive, and the availability of new frequencies is not always guaranteed : this might be a limiting factor to the development of wireless networks in some geographic areas. As a consequence, a large part of the existing literature deals with the performance of the physical layer and investigates new fashions to improve the efficiency of the communication medium. The most widespread solutions to improve the AMI physical layer performance are (i) the adoption of new coding schemes (e.g., in Arias and Rodriguez (2015); Ayala and Rodriguez (2014); Le and Benjapolakul (2016)), (ii) the exploitation of new spectrum opportunities (e.g., in Riascos et al. (2015), Parvez et al. (2016), Khan et al. (2016), and Sousa et al. (2015)), and (iii) the employment of new modulation techniques (e.g., in Oh et al. (2015) and Zeng et al. (2012)).

3.3.4 Routing and network layer

AMIs are conceived to connect the power utility MDMS to all smart meters in its power grid. Managing the routing of packets in such a large and heterogeneous network is not straightforward. Sabbah et al. (2014) provides an interesting overview of the networking difficulties in the context of Smart Grid. First, it is difficult to design a routing protocol able to work with the heterogeneous traffic produced by the diversity of Smart Grid applications. Then, a wide variety of different technologies coexist in a Smart Grid infrastructure : accordingly, a multitude of different routing protocols — each suitable for a subset of the technologies — have been proposed, leading to a lack of standardization and difficult inter-operability between the various segments of the architecture. In Saputro et al. (2012), a large number of routing protocols are evaluated and subdivided by the layer of AMI architecture in which

they are used, and by the types of applications for which they are well suited. Another limiting factor of routing in AMIs is that smart meters just provide basic communication features, with a limited throughput; therefore, routing protocols have to be simple and not computationally hungry : for instance, smart meters have limited storage capacity and cannot host very large routing tables nor implement complex routing algorithms. The most widespread routing mechanisms adopted in AMIs are Routing Protocol for Low-power and lossy networks (RPL), Ad-hoc On-demand Distance Vector (AODV), *geographic-based*, and *layer-based*, as discussed in Ramirez et al. (2015), Elyengui et al. (2015), and Hu et al. (2015).

RPL, thoroughly described by Wang et al. (2010) and Tripathi et al. (2010), is a routing mechanism compatible with the IPv6 standard. It is based on the use of Directed Acyclic Graphs (DAGs) to create a topology in which each node has a *rank*, to represent its position with respect to the others. DAG formation process is started by the collector and the metrics is based on the expected transmission count, computed using the acknowledgements of the Media Access Control (MAC) layer. Opportunistic RPL (ORPL) is an interesting variant of the RPL protocol : it aims at increasing AMI reliability by exploiting the wide variety of paths (i.e., DAGs) which characterizes AMI topologies, as discussed in Gormus et al. (2011).

AODV, a reactive protocol proposed in Perkins and Royer (1999), was primarily conceived for mobile ad-hoc networks, but is also used in networks whose topology frequently changes : routes are established and maintained only when necessary. Route discovery is performed with simple control messages, such as route requests, route replies, and route errors. Some variants of AODV use *Hello packets* to improve local connectivity management. The performance of the AODV protocol in the Smart Grid context is investigated in Cheng et al. (2013), Farooq and Jung (2013a), Pozveh et al. (2016), Farooq and Jung (2013b), and Kathuria et al. (2013).

Dynamic Source Routing (DSR) is a family of protocols in which the path is fully decided at the source Johnson et al. (2007). Its advantage with respect to AODV is that the complexity is located at the root node; on the other hand, AODV is more resistant to link failures and more reactive to changes in general. The so-called *layer-based* routing, discussed in Chen et al. (2012), is a DSR protocol and is widely employed in AMIs. The protocol is based on the *layer-index*, which is assigned to each node in the *layer-formation* phase and is updated by means of the *layer-updates*, small packets which are exchanged in case of route changes. According to this protocol, the data collector has index 0, its neighbors have index 1, and so on. The hierarchy is structured so that each node chooses just one neighbor (with lower layer index) and transmits to it all the uplink packets. On the other hand, the routing for the downlink packets is determined by the data collector, which is the source. This routing

protocol is simple and dynamic : it can quickly react to topology changes by using the *layer-updates* to establish new routes.

Geographical routing is also quite popular in AMIs : this family of protocols relies on GPS coordinates only. Examples can be found in Xiang et al. (2012) and Zahariadis et al. (2009). Zahariadis et al. (2009) presents a model aiming at insuring a reliable routing in very extended Wireless Sensor Networks (WSNs), while Xiang et al. (2012) achieves better performance in terms of packet loss and reliability, with an increased latency of only 2 ms.

Another well-known family of protocols used for AMI routing is *tree balancing*, similar to RPL, with a topology composed of several trees with the collector on top. Kulkarni et al. (2012a) proposed TREEB, a tree balancing method in which the state of concentrators is included within the information provided to the smart meters; a balanced repartition of flows is obtained by penalizing routes beyond a certain utilization and crowded routers.

3.4 Different approaches to the performance evaluation for AMIs

Besides the choice of different technologies, protocols, and architectures, the existing body of literature addressing the performance evaluation of AMIs can be categorized according to the main followed approaches, which are (i) field trials, (ii) mathematical analysis, and (iii) network simulation.

Field trials are used to analyze the performance of already implemented systems, allowing to catch unexpected properties (e.g., unattended blockages, or failures) which are hard to be detected with simulations or analyses. Nevertheless, the large number of devices (expensive and difficult to procure) and data confidentiality issues hinder the widespread use of field-trials for AMIs. A properly designed network simulator allows to reproduce the standard operation of an AMI in a virtual environment at reduced cost with respect to field trials. This is particularly useful in order to conceive and evaluate new system features (e.g., new techniques to mitigate wireless interference). Unfortunately, the associated computational burden may hamper the applicability of this methodology to very large-scale AMIs. Finally, the analytic approach permits to reduce the computational cost, with respect to simulations. Moreover, the outcome of mathematical analyses can easily be extended to different scenarios and case studies. Nonetheless, in order to reproduce complex system details in a mathematical model, it is necessary to use assumptions that might undermine the model accuracy and fitness to reality.

Additional details on the analytic and field-trials approaches — especially centered on RF-Mesh AMIs — can be found in Sections 4.2 and 5.2; while a thourough analysis of the previous simulations studies of AMIs is reported in 6.1.

CHAPTER 4 ARTICLE 1 : ANALYTICAL PERFORMANCE ANALYSIS OF A LARGE-SCALE RF-MESH SMART METER COMMUNICATION SYSTEM

F. Malandra et B. Sansò, (2015), "Analytical performance analysis of a large-scale RF-mesh smart meter communication system", Innovative Smart Grid Technologies Conference (ISGT), 2015 IEEE Power and Energy Society, Washington, DC, 2015, pp. 1-5. `http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7131840&isnumber=7131775`.

Abstract

Advanced meter infrastructures (AMIs) are now widespread and their importance within smart grid systems continues to increase with the advent of new applications. Performance analysis of the infrastructure is key to assess the limits of application deployment. However, due to the large-scale nature of AMI networks that are often composed of tens of thousands of nodes per collector, performance analysis is often carried out in contained experimental trials. To our knowledge, no thorough mathematical performance analysis of real-sized systems has been carried out so far. In this work, we present a model to analyze the performance of a large-scale RF-AMI system and show its application to large-scale real-case scenarios.

4.1 Introduction

AMI systems are increasingly popular and massive worldwide installations of smart meter devices is currently ongoing. Several studies forecast an acceleration of this process in the following years, mainly driven by the large revenues expected from the use of smart meter applications within the smart grid market.

Several technologies can be adopted but the RF-mesh based system seems to be one of the most popular. RF-mesh systems are mainly used for remote reading, advanced metering and for some other applications, such as demand-response or load management, that do not have strong requirements in terms of bandwidth and delay. However, utilities, that have spent millions of dollars installing such a widespread communication infrastructure, may want to exploit it for other types of applications that, in some cases, may require shorter response time.

A fundamental question that arises in this context is how to be able to assess the limits of the installed infrastructure. Two methods can be though of : stochastic simulation and field

trials. The difficulty of stochastic simulation lays on the size of the system, that may contain thousands of nodes, thus requiring specialized codes for parallel implementation and still would need a very large amount of resolution time Lichtensteiger et al. (2010). On the other hand, field trials are limited, because they depend not only on the particular conditions of the households involved in the trial, but also on the time those trials are carried out. In fact, depending on the application, load conditions may influence the amount of communication that can be exchanged and load conditions greatly vary with time of day, month and season, among others. Moreover, the nature of field trials is observation-based, as, most of the time "what ifs" cannot be easily implemented.

There is, therefore, an important need for a flexible analytical model for large scale RF-smart meter performance evaluation that would allow not only to assess the current system, but also do extended "what ifs" studies, study the suitability of the system for future applications as well as help in the assessment of the features most needed in the evolution of the system. The object of this work is precisely to fill this gap by proposing an analytical characterization of large-scale RF-smart meter performance.

The document is structured as follows : section 4.2 is a brief overview on the literature concerning the performance of wireless mesh networks, with a particular attention to smart grid applications; section 4.3 presents the modelling of the system as well as some relevant performance indexes; section 4.4 reports some results; section 4.5 summarizes the conclusions of this work.

4.2 State of the art

An RF-based AMI system is a largely distributed wireless mesh network for which it is extremely difficult to put in place stochastic simulation results given its large scale. Nevertheless, some authors have attempted this approach (e.g., AlMajali et al. (2012); Iyer et al. (2011); Lichtensteiger et al. (2010); Patel et al. (2011)). Others have dealt with real-time trial measurements. For instance, Cespedes et al. (2012) reported some results obtained in a real AMI while Kulkarni et al. (2012b) combined the simulation approach with real-time measurements.

Our methodology is totally analytic and, in that respect, it is based on fundamental findings of wireless networks performance.

Interference among the links is a key issue to assess wireless mesh network performance. Jain et al. (2005) defined a *connectivity graph* and associated to it a *conflict graph*, using the information on nodes, links and distances. The problem of maximum throughput was

transformed into the search of the maximum independent set (a set of vertices that can transmit simultaneously), which is a NP-hard problem and which does not guarantee a solution. Nandagopal et al. defined a *contention* graph using flows rather than links to account for interference : the objective of the authors was to provide a *fair* access to the medium to different commodities. Kodialam and Nandagopal (2003) tackled the problem of performance analysis in a multi-hop wireless network focusing on routing and scheduling. The authors modelled the problem as a graph-coloring one and assumed a simplified interference model in which each node cannot transmit and receive on the same channel at the same time. In our work, we took inspiration from this paper in what concerns the interference analysis, that is one of the steps needed to develop the analytic tool for performance evaluation.

The other issue that has a strong impact in the overall network performance is the effect of the MAC layer, that, in the case under study is ALOHA based. Performance of the ALOHA system has been well documented (e.g., Kaynia and Jindal (2008); Médard and Goldsmith). Our system differs from the other in literature because of the large number of nodes ($\approx 10^3$) and of the use of Frequency Hopping Spread Spectrum (FHSS) protocol to tackle interference (see section 4.3 for further details).

Even though there has been some simulation and field studies on AMI performance as well as a rich literature on mesh network performance, to the best of our knowledge this is the first time that a comprehensive analytical model is put in place to assess several performance parameters of a large-scale AMI system. Our analytical framework is particularly useful because it can be adapted to different types of scenarios and network features and it does not present the computational issues of very large scale systems simulation.

4.3 RF mesh communication system modelling

4.3.1 Main features

Many deployed AMI systems present proprietary features that are not easily disseminated to the public. So, in this paper, the case-study was a deployed system for which some data is publicly available. The communication system in object is a multi-hop wireless mesh network made of smart meters, routers and collectors with a current infrastructure of 1.5 million smart meters. The technology to connect the meters to data collectors is RF mesh and routers are used to extend coverage and increase connectivity. The communication channel is the unlicensed ISM band of $902 - 928$ MHz. The system uses FHSS, which is a technique particularly efficient against low spectrum interference coming from other devices transmitting on the same free band. The access to the medium is regulated by a synchronous

ALOHA with time slots of duration $\tau = 0.7$ s. Devices are kept synchronized through the Network Time Protocol (NTP). Collectors are the only devices equipped with GPS receivers, routers and smart meters are synchronized to their collector with NTP with an offset in the order of some milliseconds.

4.3.2 Topology definition

The topology definition in a large scale metering system is a challenging task. For this particular case, we used public information to extract the general areas where smart meters were installed, that happened to present the position of routers and collectors. Next, we calculated the GPS coordinates by means of Google Maps and Bing Maps application programming interfaces. To locate the smart meters, we assumed one device per home and we developed a script in Bing Maps that inputs public data, such as postal codes or list of municipality streets, and outputs the latitude and longitude of the points. At the end of this process, as shown in Figure 4.1, we obtained the position of the nodes of the network. To define links, two different covering rays were assumed : 0.15 to 0.5 km for smart meters and of up to 2 km for routers and collectors. The variation of the maximum transmission range of smart meters is due to different propagation conditions in urban, rural and suburban areas. Moreover, routers and collectors achieve extended coverage with directive antennas and top-building installation. Routers and collectors have higher capacity links (19.2 kbps) with respect to smart meters (9.6 kbps), as shown in Figure 4.1.

4.3.3 Shortest paths

To start the analysis, a static shortest path routing based on distance as a metric was assumed but later on, other types of routing will be considered. The problem can be seen as a large multi-commodity flow since different streams of communication coexist in the system. There is one commodity from each smart meter to the collector and vice-versa.

4.3.4 Traffic characterization

Let $G(V, E)$ characterize the graph representing the AMI system, where V and E are the sets of nodes and links of the RF network. The set of nodes V is composed by the union of the set of smart meters M, the set of routers R and the set of collectors C.

Let λ_{up} be the mean traffic from each smart meter to the collector and λ_{down} the mean traffic from the collector to a single smart meter. In both cases, Poisson arrivals are assumed so

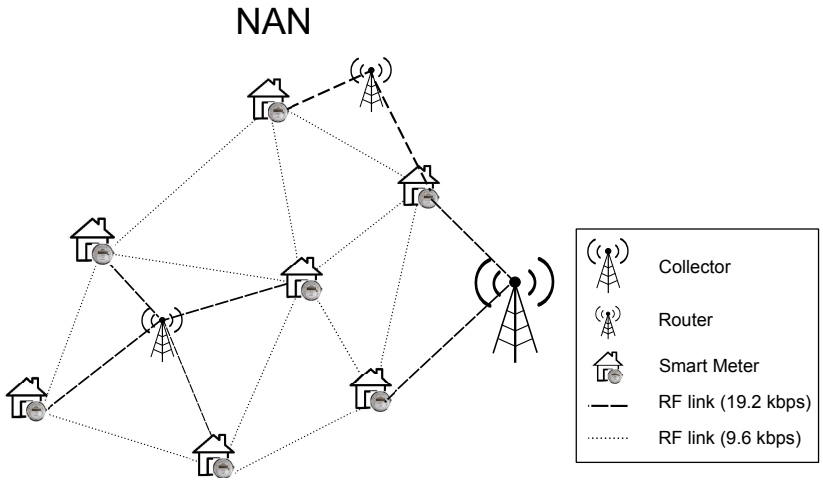


Figure 4.1 Topology definition.

that the aggregation of traffic streams at each node is also Poisson with a mean value equal to the sum of all sub-streams mean values.

Let λ_i be the transmission rate of generic device i. In order to characterize λ_i,the routing behaviour of each device must be analyzed. Intermediate nodes are in charge of transmitting packets from the origin to the destination of the shortest paths they belong to. For this purpose, we introduce ξ_j, the number of shortest paths that contain node j as an intermediate node. Then, a given smart meter j transmits its own packets to the collector at a rate λ_{up}, packets from ξ_j smart meters to the collector at a rate $\xi_j\lambda_{up}$ and finally packets from the collector to the ξ_j nodes at a rate $\xi_j\lambda_{down}$. Routers do not generate traffic so a given device i is in charge of transmitting packets of ξ_j different streams in both uplink and downlink directions. On the other hand, collectors transmit to each of the $|M|$ smart meters with a rate λ_{down}.

These considerations can be summarized as follows :

$$\lambda_i = \begin{cases} \xi_i(\lambda_{up} + \lambda_{down}) + \lambda_{up}, & \text{if } i \text{ is a smart meter} \\ \xi_i(\lambda_{up} + \lambda_{down}), & \text{if } i \text{ is a router} \\ |M|\lambda_{down}, & \text{if } i \text{ is a collector} \end{cases} \tag{4.1}$$

The actual transmission rate also depends on the number of packet retransmissions that are caused by collisions. In particular, if N_i represents the average number of retransmissions

at node i, the actual transmission rate can be defined as $\tilde{\lambda_i} = N_i\lambda_i$. The average number of retransmissions and the link with the collision probability is deeply analyzed in section 4.3.5.

4.3.5 Probability of collision

One of the main issues to a wireless communication system is interference. If we assume that the interfering ray equals the covering ray, every node's neighbours are its possible interferers. If we neglect for the moment FHSS, we can state that there is a collision at node i when at least one of its neighbours attempts to transmit during the same time slot. When a collision is experienced, the involved packets have to be retransmitted. The average number of times a packet at node i is retransmitted, N_i, is related to the average collision probability by :

$$N_i = \frac{1}{1 - p_i} \tag{4.2}$$

As stated in the previous section, the transmission rate of a set of nodes has a Poisson distribution with the sum of all λ_i as mean value. Given I_i the set of interferers of node i, the probability that none of the nodes $j \in I_i$ transmits in a time-slot of duration τ is :

$$P(X_I = 0) = e^{-\tau \sum\limits_{j \in I_i} \tilde{\lambda_j}} = e^{-\tau \sum\limits_{j \in I_i} \lambda_j N_j} = e^{-\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}} \tag{4.3}$$

where X_I is the number of nodes in I_i that transmit during that time-slot. Then, the probability that collisions occur is :

$$p_i = P(X_I > 0) = 1 - P(X_I = 0) = 1 - e^{-\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}} \tag{4.4}$$

If we want to take into account FHSS, the definition of collision changes : a collision occurs at node i when at least one of its neighbours is transmitting on the same channel as i. In our first FHSS modelling, we consider that each node randomly decides the transmission channel among the Q available. Let us consider the event that at least one of the k neighbours of i chooses the same channel as i. This is the complement of the event in which all the k nodes choose different channels than i; therefore, the probability that at least one of the k nodes chooses the same transmission channel as i is given by :

$$p^{(k)} = 1 - \left(1 - \frac{1}{Q}\right)^k \tag{4.5}$$

As a consequence, a collision occurs when there is at least one transmitting neighbour of node i using the same channel as i is calculated as follows :

$$\begin{aligned} p_i &= \sum_{g=1}^{+\infty} P(X_I = g) p^{(g)} = \\ &= \sum_{g=1}^{+\infty} \left(1 - \left(1 - \frac{1}{Q}\right)^g\right) \frac{\left(\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}\right)^g}{g!} e^{-\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}} \end{aligned} \tag{4.6}$$

This is a so-called *fixed point equation*. The problem of finding the values of p_i (for $i = 1 \dots |V|$) that solve the non-linear system (4.6) is equivalent to solving the following optimization model :

$$\min_{\boldsymbol{p}} \|f(\boldsymbol{p})\|_2^2 \tag{4.7}$$

$$s.t. : \begin{cases} p_i \geq 0 & \forall p_i \in \boldsymbol{p} \\ p_i < 1 & \forall p_i \in \boldsymbol{p} \end{cases} \tag{4.8}$$

where :

$$\boldsymbol{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{|V|} \end{bmatrix} \quad f(\boldsymbol{p}) = \begin{bmatrix} f_1(\boldsymbol{p}) \\ f_2(\boldsymbol{p}) \\ \vdots \\ f_{|V|}(\boldsymbol{p}) \end{bmatrix} \tag{4.9}$$

and $f_i(\boldsymbol{p})$ is given by :

$$f_i(\boldsymbol{p}) = \sum_{g=1}^{+\infty} \left(1 - \left(1 - \frac{1}{Q}\right)^g\right) \frac{\left(\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}\right)^g}{g!} e^{-\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}} - p_i \tag{4.10}$$

One can note that, when the number of channels is 1, (4.6) is equivalent to (4.4) since $\sum\limits_{n=1}^{+\infty} \frac{a^n}{n!} = e^a - 1$.

The performance analysis carried out in the rest of this work is based on the vector $\boldsymbol{p}$, solution of the optimization model (4.7). At this time, numerical results were obtained using the simplified equation (4.4), neglecting in the first phase the presence of FHSS.

4.3.6 Delay

A basic parameter in the analysis of the performance of a telecommunication network is delay. In a multi-hop random access system with small packets, such as Time-Slotted Aloha, it is a standard practice to let the time slot duration include all other types of delay a packet can encounter. In the rest of the discussion, propagation, transmission and processing delays are included in the $0.7s$ time slot. In this first analysis, the queueing delay is considered negligible since we are dealing with very few traffic rates.

In the light of these considerations, we calculate the average delay based on the number of hops a packet makes in order to reach its destination only. The delay of a generic n-hop path such as the one displayed in Figure 4.2 can be calculated as follows :

$$d_{ij} = \tau \sum_{k=0}^{n-1} N_k \tag{4.11}$$

Note that in (4.11), delay depends exclusively on the length of the shortest path as well as the probability of retransmission of the system. If queueing and processing delays were not negligible, (4.11) should be modified by adding the appropriate M/M/1 delay modelling terms.

We account for two different types of delay :

d_i^u *uplink delay*, the time necessary for a packet generated by a smart meter to get to the collector;

d_i^d *downlink delay*, the elapsed time for a packet to travel from the collector to a smart meter.

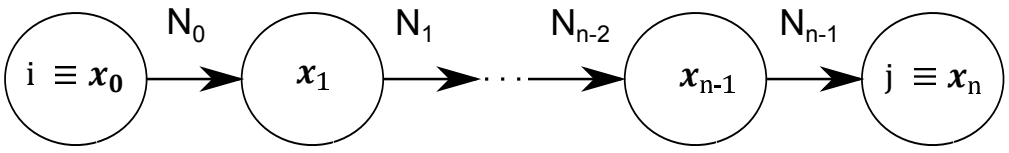


Figure 4.2 Scheme of the n-hop path from node i to node j.

4.4 Results

This section presents some numerical results obtained with the model introduced in the previous sections. We have chosen a typical rural area with an extension of $240km^2$ with 3415 nodes (1 collector, 114 routers and 3300 smart meters).

Moreover we defined 32 different traffic scenarios, based on message exchange rate on the uplink, λ_{up}, and the downlink, λ_{down}, assuming that all the meters had the same message frequency (see Table 4.1). The reader should be aware that such an assumption was adopted to ease the interpretation of results, but it could be easily relaxed within our modelling framework.

Results were obtained using MATLAB, with an Intel(R) quad Core(TM) i7-3770 CPU @ 3.40GHz processor. The average computational time was 1684.86 s with the maximum value of 7250 s obtained in the 3^{rd} scenario and the minimum of 342.52 s in the 32^{nd}.

4.4.1 Collision probability

The solution of the set of equations (4.4) provided the probabilities of collision that each node is subject to when it attempts to transmit a packet. Figure 4.3 represents some statistics of the collision probability in different scenarios : the maximum and the average values related to each kind of device are reported. The black curves with triangles are related to routers. The maximum is, in many scenarios, close to 1 while the average goes from 0.35 in the first scenario to less than 0.1 in the last scenario. The grey curves with crosses refer to the maximum (continuous line) and the average (dotted line) values related to the smart meters. In this case, both maximum and average collision probabilities are considerably lower than those of routers. We can also remark that the average collision probability for smart meters is in all scenarios below 10%. When it comes to the collector, we can see that there is only one curve since we have a unique device of this type in each scenario. The dashed-dotted

Table 4.1 Traffic scenario ID according to the values of λ_{up} (columns) and λ_{down} (rows).

$\frac{1}{\lambda_{down}}$ [h]	$\frac{1}{\lambda_{up}}$ [h]							
	0.5	**1**	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**
1	1	5	9	13	17	21	25	29
2	2	6	10	14	18	22	26	30
3	3	7	11	15	19	23	27	31
4	4	8	12	16	20	24	28	32

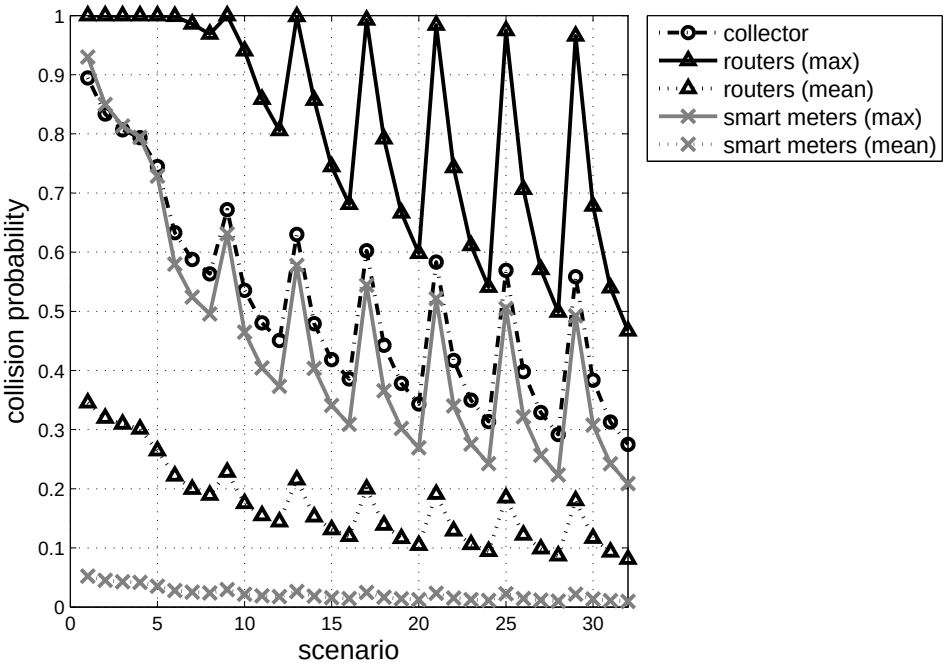


Figure 4.3 Statistics of the probability of collision of different devices in different scenarios.

black line with circles shows that the performance of the collector is in between that of smart meters and routers : its collision probability is not excessively high (it is above 7% only in the five first scenarios).

4.4.2 Delay profile

Provided we calculated the delay as the sum of the number of retransmissions in each hop (4.3.6) and that the average number of retransmissions is inversely proportional to the probability of collision (4.2), we do not show the trend of delays in the system since it is very similar to that for the probability of collision presented in the previous section.

4.4.3 Survival function : application feasibility

A interesting insight on the performance of our system can be brought by the notion of the percentage of smart-meter nodes presenting an uplink or downlink delay above a certain threshold. In the smart grid context, it is important that the received information be up-to-date, and this is true only if it is received within a short delay.

In order to carry on this analysis, we use the concept of survival function. Let Y be a random variable, its survival function is $f(y) = P(Y > y)$ with $y \in [0, \infty]$.

Let D_u and D_d be the discrete variables related to uplink and downlink delay respectively, we define two survival functions as follows :

$$f_u(\gamma) = P(D_u > \gamma) \tag{4.12}$$

$$f_d(\gamma) = P(D_d > \gamma) \tag{4.13}$$

The notion of survival function allows us to make conclusions about the feasibility of smart grid applications in the RF-mesh system.

Let α and β be two generic load management applications with the following requirements : the collector transmits a packet every 4 hours in average, each smart meter transmits a packet every 2 hours in average; a packet is considered old by the collector if received after a delay (in uplink) of 20 s. α can tolerate a maximum of 20% of packets to be old, whereas β can tolerate only 10%. The traffic transmission rates suggest that our example corresponds to scenario 16 ($\frac{1}{\lambda_{down}} = 4$ h, $\frac{1}{\lambda_{up}} = 2$ h). We are analyzing the delay at the collector side, that is to say the uplink delay.

Figure 4.4 shows the survival function of delay in uplink in scenario number 16. In the same graph we plot the points (20 s, 0.2) and (20 s, 0.1), relative to the requirements for α and β. The first coordinate of these points represents the delay threshold (γ), while the second represents the maximum tolerated probability $P(D_u > \gamma)$. If the point is above the curve, as for α, it means that the requirements of its associated application are within the possibilities of the system : the application is considered feasible. On the other hand, if the point is below the curve, as for β, it means that the associated application requirements cannot be met by our system : the application is not feasible.

4.5 Conclusions

This work presents, to our knowledge, a first analytical model for the overall performance evaluation of an AMI RF-mesh based system. The model becomes an important tool since it bypasses the difficulty of having to perform costly stochastic simulations for a network containing thousands of nodes, while enhancing the quality of information that can be extracted from field trials. The tool allows different types of what-if studies to assess the applicability of forthcoming applications, which is quite important for utilities.

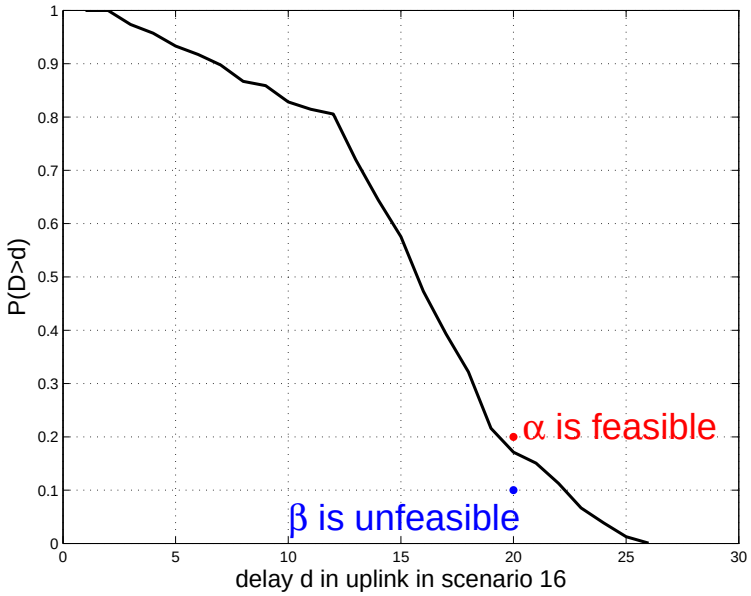


Figure 4.4 Survival function of uplink delay in scenario 16.

We have not just developed a model, but also defined a framework to assess AMI RF-mesh performance. This framework is based on the type of routing, the interference model, the traffic analysis and the type of measures that must be taken into account to assess network performance, such as delay, collision probability, critical nodes and a survival function particularly useful to assess the feasibility of the deployment of new applications.

We constructed a study case extracting publicly available information. For such a case, even though the conclusions were expected, it was possible to quantified them with the tool. The traffic rate (λ_{up} and λ_{down}) considerably affected the performance of the network and the shortest path routing was not very efficient. The model, however, can easily incorporate any other type of routing. Also expected was the fact that routers are the critical elements of the system, as the percentage of routers in critical conditions is much higher than for the other devices.

Finally, it is clear that in order to carry out real-time or quasi-real-time applications, the current infrastructure needs to be upgraded or must co-exist with some off-loading network such as the Internet. We are currently investigating performance and optimization models to take into account those issues.

Acknowledgments

This work was supported by an ECO Energy Innovation Initiative grant from Natural Resources Canada.

CHAPTER 5 ARTICLE 2: PERF-MESH : A PERFORMANCE ANALYSIS TOOL FOR LARGE SCALE RF-MESH-BASED SMART METER NETWORKS WITH FHSS

F. Malandra et B. Sansò, (2015), "PeRF-Mesh : A Performance Analysis Tool for Large Scale RF-mesh-based Smart Meter Networks with FHSS", 2015 IEEE International Conference on Smart Grid Communications (SmartGridComm), Miami, FL, 2015, pp. 792-797, `http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7436398&isnumber=7436263`.

Abstract

This work deals with the performance analysis of a particular type of AMI : the RF-mesh based smart meter network. The system implements a MAC access with a time-slotted ALOHA with the Frequency Hopping Spread Spectrum (FHSS) to reduce co-channel interference by other users. We developed the PeRF-mesh analytic tool to study the performance of such systems, taking into account the combined effects of ALOHA access and of FHSS on the performance. The tool allows the evaluation of currently deployed systems and can also help in the design phase of new ones.

5.1 Introduction

In many countries around the world, power utilities have already equipped a large percentage of households with smart meters; others are planning on a comprehensive installation process in the forthcoming future. Smart meters assume a key role in many smart grid applications because of their double nature of sensing and communicating devices. In order to accomplish their functions, smart meters need to have a two-way communication link with the power utility management system : this is the main reason why the penetration of Advanced Metering Infrastructure (AMI) is very deep within smart grid systems.

AMIs are large scale systems in which thousands of nodes are involved (e.g. sensors, smart meters, routers, data collectors) and many applications are enabled (e.g. remote reading, load management and Vehicle-to-Grid). They are usually proprietary systems, owned by power utilities and installed by third party companies.

Several technologies have been adopted and installed for AMI so far : some are based on the use of the Internet, employing different types of access (mainly cellular or WiFi), while others exploit the presence of electric wires by using Power Line Communication (PLC). Further

solutions consider the use of radio frequencies in free and unlicensed bands : for example, for RF-mesh, the Industrial, Scientific and Medical bandwidth from 902 to 928 MHz is used.

RF-mesh is considered one of the most popular technologies within AMI systems and it will be presented with further details in Section 5.3. It is characterized by a simple architecture composed of smart meters, routers and data collectors; RF antennas are cheap and the infrastructure is proprietary, feature researched by the power utilities that do not want to rely on telecommunication providers, mainly because of cost and data confidentiality reasons. Nevertheless, some of the advantages of this technology can also be seen as shortcomings : the absence of a recognized standard within the plurality of vendors jeopardizes the interaction of one system with another; also, it is difficult to define the performance of proprietary systems because many features of the devices are covered by confidentiality agreements. Moreover, a very low data-rate is achievable with this technology : the nominal throughput is in the order of tens of kilo-bits, a number that sounds anachronistic but which can still enable many smart grid applications.

The multitude and vicinity of nodes, which especially characterize urban environments, can lead to severe interference problems. The issue is tackled by employing Frequency Hopping Spread Spectrum (FHSS) protocol. To the best of our knowledge, none of the existing comprehensive analytic studies on the performance of large scale mesh systems considers the effect of FHSS protocol. However, as showed in Section 5.5, where we compare numerical results with and without FHSS, this protocol has a fundamental importance in large scale RF-mesh systems.

RF-mesh systems are usually sold as *black boxes* to power utilities and many questions arise when it comes to the analysis of the performance : what is the average delay? How many routers are necessary to cover a given area? How many packets can be received on time, coping with peculiar applications requirements?

The objective of our work is to provide answers to the aforementioned questions by means of PeRF-mesh, the analytic tool we implemented, helpful in defining measures and indexes of performance for large scale RF-mesh based smart meter communication systems.

The document at hand is structured as follows : Section 5.2 contains a short literature review centered on the performance analysis in wireless mesh networks, with a particular focus on smart grid systems; Section 5.3 presents the modeling of the system under consideration; Section 5.4 describes PeRF-Mesh, the analytic tool for performance evaluation; in Section 5.5 some numerical results are shown and in Section 6.5 the conclusions of the present work are summarized.

5.2 State of the Art

A great deal of research effort is currently being expended on the performance study of RF-mesh networks. The importance of this theme is derived from the increasing interest in new smart grid applications. The main approaches that have been followed in literature can be grouped in two sub-categories : stochastic simulations AlMajali et al. (2012); Iyer et al. (2011); Lichtensteiger et al. (2010); Patel et al. (2011) and real-field measurements Cespedes et al. (2012); Kulkarni et al. (2012b) .

Both approaches have some strong points as well as some shortcomings. As a matter of fact, a well configured simulator can perform significant performance studies with great savings, and can be helpful in designing and testing new and not yet implemented features and solutions for existing systems. On the other hand, real-field measurements permit analyses of actual systems and not of a modeled version of these. Also, testing systems in a real environment can give deeper insights on their characteristics : some features (e.g. realistic propagating conditions) are very difficult to predict and model, and real field tests can cast light on inconsistencies of the model, which a simulator could hardly discover because of the ideal environment it works within.

A third approach, which we decided to follow, is totally analytic : known properties of wireless networks are used to find mathematical equations that allow to analyze the system's performance. The analytic methodology can reduce the computational burden typical of simulations and can be easily extended to different scenarios and technologies.

The wireless interference problem is well explained in Li (2008), where several protocols to model interference are presented. One of the most used, which we chose to adopt, is the *protocol-interference model*. In this model, first presented in Gupta and Kumar (2000), all the nodes at a certain distance from a node j are considered possible interferers in a communication directed to node j.

In a large scale network with thousands of users that share the same bandwidth, the performance is clearly affected by the choice of the MAC layer protocol : one of the most widespread in RF-mesh systems is the slotted ALOHA. An extensive research focusing on ALOHA performance has been carried out since its first presentation in 1971 by Norman Abramson Abramson (1970). The pioneer works of Carleial and Hellman (1975); Gitman (1975); Kleinrock and Lam (1975) laid the foundations of the analytic performance study of slotted ALOHA systems, focusing on single-hop systems only. Silvester and Kleinrock (1983) tried to analyze multi-hop systems with simple and regular topology (e.g. loop and bus). We took inspiration from the extensive work on ALOHA performance analysis models in order

to find a mathematical equation for the collision probability in a RF-mesh system. To the best of our knowledge, a comprehensive analytic study of the combined effect of ALOHA and FHSS protocols on network performance is not available in literature.

5.3 RF-mesh system architecture and main features

One of the main difficulties in modeling AMIs is the fact that these are proprietary systems and many of their features are undisclosed. For this work, features of the RF-mesh smart meter communication network are derived from publicly available data about a RF-mesh system already installed in Québec Hydro-Québec (2012).

The system under study has a three-layers architecture, as shown in Figure 5.1 : the first is the Home Area Network (HAN), that consists of sensors, smart meters, appliances and all the other devices within the domestic area; the second is the Neighborhood Area Network (NAN), whose main scope is to connect smart meters (and consequently the HAN) to data collectors in a mesh topology that also includes routers; data collectors are used as gateway to the third layer of the architecture, the Wide Area Network, an IP backbone connected to the power utility Metering Data Management System (MDMS).

The first and the third layers of the architecture implement well-known technologies and protocols : the HAN adopts Zigbee short range links, while the WAN uses IP over satellite or cellular connections. On the other hand, the NAN is characterized by wireless links in the ISM band of $902 - 928$ MHz : this technology is called RF-mesh. The performance of the HAN and the WAN are well defined and a good branch of research involves analysis of Zigbee, satellite or cellular networks. Thus, in the rest of this work, we will focus on analyzing the performance of the RF-mesh NAN, not yet well defined and standardized.

In the currently deployed multi-hop wireless NANs, there is one data collector per several thousands of smart meters. The number of routers depends on the scenario : it is higher in rural networks with respect to urban environments, in order to ensure the connectivity in a more extended area.

The RF-mesh system adopts the FHSS protocol, which is a technique helpful in reducing co-channel interference, generated by the transmission of multiple devices using the same frequency band, either within the same NAN or in different networks [1].

The frequency spectrum of the RF-mesh system under study is subdivided in $n = 80$ channels of 300 kHz bandwidth each. A predetermined sequence of hops, schematized in Figure 5.2, is

[1] The ISM band are used, among the others, by some microwave ovens, car key remote controllers, ZigBee and RFID

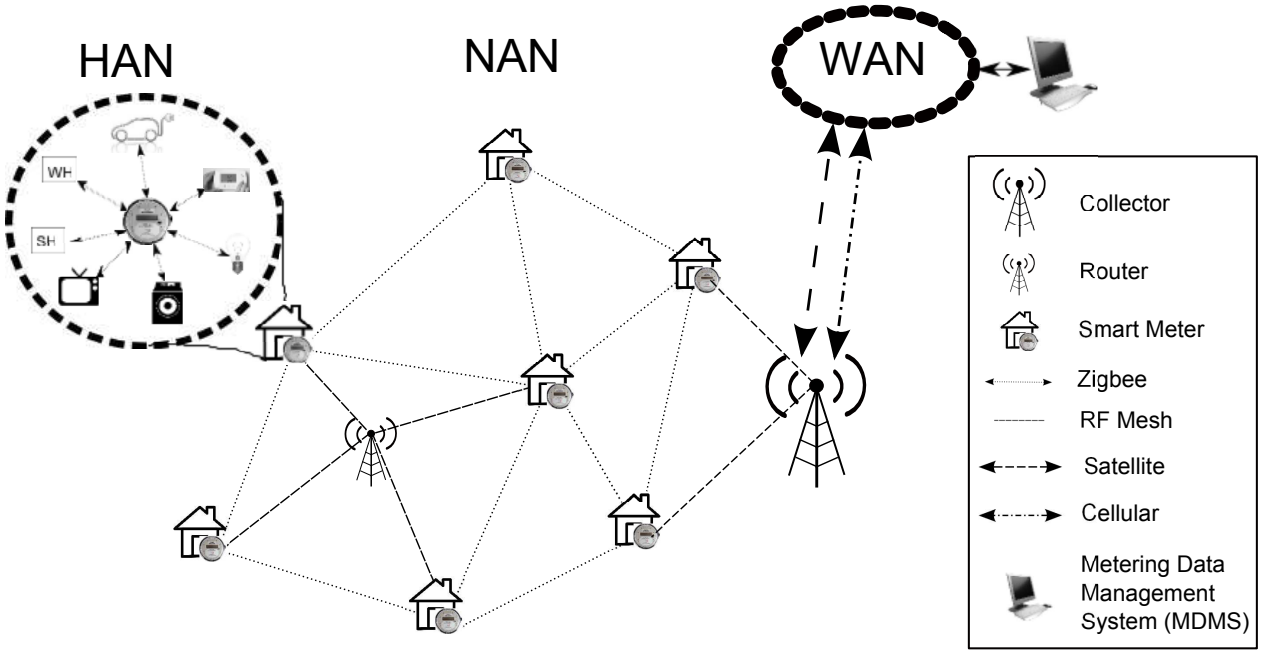


Figure 5.1 Architecture of the whole communication system.

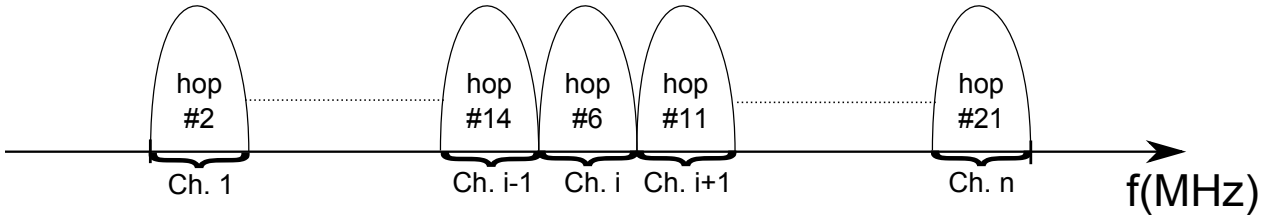


Figure 5.2 Example of frequency hopping sequence.

known to all the nodes in the network. Each device uses the same sequence to determine the frequency channel which its receiving antenna must be tuned to : the sequence is conveniently shifted in time in order to avoid that all the devices use the same channels simultaneously. Any node is able to determine the receiving frequency channel of its neighbors at any time; therefore, before transmitting a packet to a neighbor node j, node i can tune its antenna to the frequency channel of node j.

The access to the medium is controlled by means of the synchronous ALOHA protocol with time slots of duration $\tau = 0.7$ s. Devices synchronization is achieved through the Network Time Protocol (NTP) : collectors are equipped with high precision clocks (e.g. iridium) and provide a reference time for the other nodes. NTP can ideally yield good results in terms of synchronization of extended networks : unavoidable errors in synchronization are tackled by restricting the portion of time in which it is possible to transmit to only 400 ms out of the available 700 ms, thus leaving the remaining 300 ms intentionally idle as a safety margin.

5.3.1 Interference and probability of collision

Interference is one of the main limits of wireless communications : the radio channel is shared among multiple users that can interfere with each other. Therefore, any wireless technology has to consider interference and reduce its effect on performance.

In RF-mesh systems, the access to the medium is regulated by ALOHA, a simple random access protocol conceived for networks with very low data-rates. When two or more interfering users attempt to transmit a packet, a collision is experienced and the involved packets are to be re-transmitted. An analytic expression to calculate the probability of collision in a multi-hop system, valid when a Poisson distribution of traffic generation is assumed, was presented in Malandra and Sansò (2015) :

$$p_i = P(X_{I_i} > 0) = 1 - P(X_{I_i} = 0) = 1 - e^{-\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}} \tag{5.1}$$

where I_i is the list of interferers of node i, X_I is the number of transmitting nodes in set I, λ_i the mean transmission rate of node i and τ the time slot duration.

The numerical results in Malandra and Sansò (2015) were obtained using equation (5.1), without considering FHSS. As discussed in that paper, the results highlighted the necessity of integrating the FHSS protocol in the performance analysis of large scale RF-mesh systems.

A first step in the integration of FHSS was taken in Malandra and Sansò (2015) with the analytic formula :

$$p_i = \sum_{g=1}^{+\infty} P(X_{I_i} = g) p^{(g)} =$$

$$= \sum_{g=1}^{+\infty} \left(1 - \left(1 - \frac{1}{Q}\right)^g\right) \frac{\left(\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}\right)^g}{g!} e^{-\tau \sum\limits_{j \in I_i} \frac{\lambda_j}{1-p_j}} \tag{5.2}$$

In equation (5.2), Q is the number of non-overlapping channels used by FHSS and $p^{(k)}$ is the probability of having at least two nodes out of k using the same frequency channel among the Q available :

$$p^{(k)} = 1 - \left(1 - \frac{1}{Q}\right)^k \tag{5.3}$$

Equation (5.2) is used in PeRF-mesh tool to calculate the delay and define other important performance indexes, as explained in Section 5.4.

5.4 PeRF-mesh

5.4.1 Inputs

PeRF-mesh is the analytic tool we developed to analyze the performance of a large scale RF-mesh system with FHSS. Its structure is displayed in Figure 5.3.

The tool needs the preliminary definition of some inputs : topology, routing and traffic.

The characterization of the topology consists of two phases : nodes placement and links definition. We decided to create our topology starting from publicly available data about a pilot installation of smart meters in Québec. Data about the position of routers and collectors were extracted by means of Google Earth, starting from a map published in a report to the *Régie de l'énergie*[2] of Québec Hydro-Québec (2012). Since the number of smart meters in the reference topology is greater than 3000, it was not possible to acquire information about their positions from the map included in the report. Therefore, in order to find their position, we developed a Python script to obtain from *Bing Maps* the GPS coordinates of residential buildings present in the pilot project areas. The assumption of one smart meter per building was used : this assumption works well in rural areas where there is a majority of one-family buildings; in urban areas, the assumption needs to be modified in order to account

[2] An economic regulation agency of the energy market.

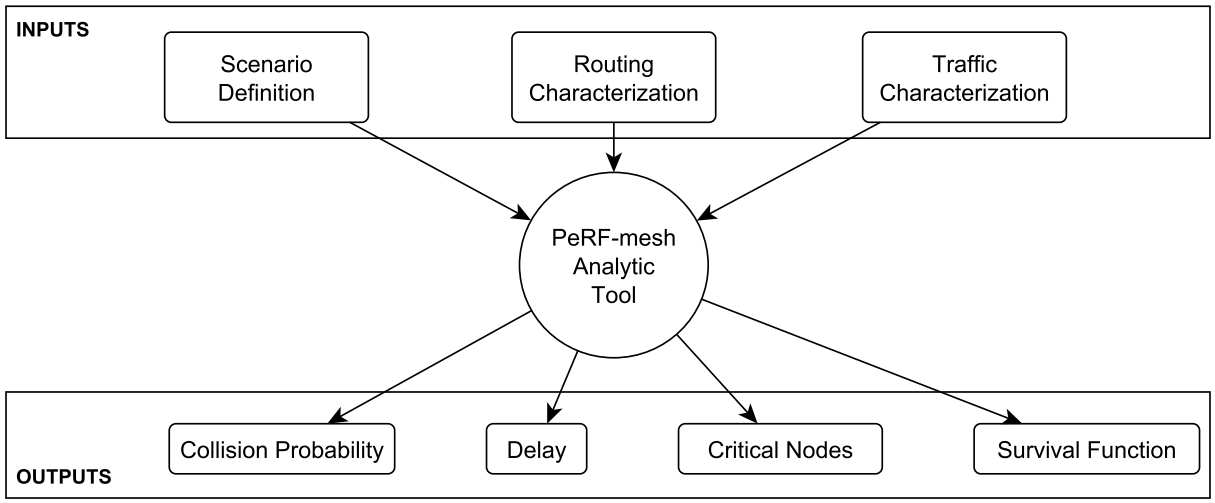


Figure 5.3 Block diagram of PeRF-mesh analytic tool.

for buildings with many apartments. The links were defined in a static way : two nodes are assumed to communicate with each other if and only if their distance is lower than a fixed covering ray. The variable propagating conditions of radio signals are taken into account by employing different covering rays in different scenarios. For example, propagating conditions tend to be more convenient in rural areas with respect to urban environments, because less obstacles are present on average; when studying a rural area scenario, larger covering rays will be used.

The routing mechanism adopted in the tool is based on shortest paths, using distance as metrics. Nevertheless, this static assumption might neglect some important dynamic aspects of RF-mesh systems : other routing mechanisms (e.g. layer-based, AODV, geographical) are currently being investigated and will be integrated into the tool in the near future.

The traffic characterization is taken from Hydro-Québec (2012). We consider two different traffic streams : uplink, from smart meters to the data collector, and downlink, in the opposite direction. Routers do not generate any packet, they simply forward packets transmitted by other devices. The packet generation rate is assumed to be Poisson-distributed in both directions with mean parameters λ_{up} and λ_{down} for uplink and downlink respectively.

λ_i is the rate of packet transmission of node i : it includes packets generated by node i and also packets for which i is an intermediate node between source and destination. In Malandra and Sansò (2015) an analytic expression for λ_i was found :

$$\lambda_i = \begin{cases} \xi_i(\lambda_{up} + \lambda_{down}) + \lambda_{up}, & \text{if } i \text{ is a smart meter} \\ \xi_i(\lambda_{up} + \lambda_{down}), & \text{if } i \text{ is a router} \\ |M|\lambda_{down}, & \text{if } i \text{ is a collector} \end{cases} \tag{5.4}$$

where ξ_i is the number of shortest paths that contain node i and $|M|$ is the total number of smart meters.

5.4.2 Mathematical modeling

Once all inputs are defined, the probability of collision needs to be calculated.

For every node of the communication system, an equation (5.2) that links its collision probability to the collision probability of its neighbors can be written. The $|V|$ equations, V being the set of nodes in the network, form a *fixed-point* system of equations.

The following least-squares optimization model (Malandra and Sansò (2015)) is used to find numerical values of p_i (for $i = 1 \dots |V|$) :

$$\min_{\boldsymbol{p}} \|f(\boldsymbol{p})\|_2^2 \tag{5.5}$$

$$s.t. : \begin{cases} p_i \geq 0 & \forall p_i \in \boldsymbol{p} \\ p_i < 1 & \forall p_i \in \boldsymbol{p} \end{cases} \tag{5.6}$$

where :

$$\boldsymbol{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_{|V|} \end{bmatrix} \quad f(\boldsymbol{p}) = \begin{bmatrix} f_1(\boldsymbol{p}) \\ \vdots \\ f_{|V|}(\boldsymbol{p}) \end{bmatrix}$$

$$f_i(\boldsymbol{p}) = \sum_{g=1}^{+\infty} \left(1 - \left(1 - \frac{1}{Q}\right)^g\right) \frac{\left(\tau \sum_{j \in I_i} \frac{\lambda_j}{1-p_j}\right)^g}{g!} e^{-\tau \sum_{j \in I_i} \frac{\lambda_j}{1-p_j}} - p_i$$

It is important to remark that equations (5.1) and (5.2) are consistent with each other : in fact, (5.2) is equivalent to (5.1) when the number of available channels is one[3].

5.4.3 Delay

The delay is one of the most important parameters in a communication system. Several types of delay are present in a communication network, but a common practice in time-slotted systems with small size packets is to consider the time slot duration to prevail over propagation, processing and queuing delay. These delay components are neglected in the current model but we are evaluating the possibility to include some of them, namely the queuing delay, in a Markov-modulated model, in course of development at the time of writing.

In an ideal system with no interference, only one transmission would be required for a single hop in a path; in reality, the presence of interference entails collisions, and each collision implies a re-transmission of the packet. Therefore, the average number of time slots necessary to transmit a packet in a single hop is :

$$N_{ij} = \frac{1}{1 - p_i}$$

As a result, the overall delay in a single hop corresponds to the time slot duration τ, multiplied by the average number of re-transmissions :

$$d_{ij} = \tau \sum_{(uw) \in \rho_{ij}} N_{uw} \tag{5.7}$$

where ρ_{ij} is the set of links forming the shortest path from i to j. In this work, we consider the delay in a multi-hop path from node i to node j to be the sum of the delays in each hop. Two delay quantities, d_i^u and d_i^d, are defined, related to uplink and downlink streams of communications, respectively.

5.4.4 Other outputs

As shown in Figure 5.3, PeRF-mesh provides two additional performance indexes, previously introduced in Malandra and Sansò (2015) : the critical nodes in the system and the so-called *survival function*.

[3]This is due to the fact that $\sum_{n=1}^{+\infty} \frac{a^n}{n!} = e^a - 1$

A node is considered critical if and only if its collision probability is above a certain threshold. Such an analysis is very useful to discover eventual bottlenecks of the system.

The survival function is a mathematical function that represents the probability that a random variable is greater than a certain value. If applied to delay statistics, the survival function can provide interesting insight in the feasibility of generic smart grid applications, whose requirements are limited to a certain portion of nodes. An example of feasibility assessment using the survival function was provided in Malandra and Sansò (2015).

5.5 Numerical results

In this section, some numerical results obtained with PeRF-mesh analytic tool are presented.

We chose to test our methodology using data related to Mansonville, a rural area in Québec and one of three zones involved in the aforementioned pilot installation of smart meters Hydro-Québec (2012), in 2011. The area is extended over 240 km^2 and includes 3415 devices (1 data collector, 114 routers and 3300 smart meters).

We assumed the same packet generation rate in uplink (λ_{up}) for all the smart meters and also the same packet generation rate (λ_{down}) from the collector to every smart meter. In multiple runs, we let the mean packet generation times ($1/\lambda_{up}$ and $1/\lambda_{down}$) vary in the interval between 0.5 and 4 hours in order to highlight the performance of the system at different traffic loads, representative of different smart grid applications.

The system of equations (5.5)-(5.6) was solved by using MATLAB on a Intel(R) Quad Core(TM) $i7-3770$ CPU @ 3.40GHz processor. The average computational time was below 15 minutes.

5.5.1 Collision probability

In Figure 5.4 we reported the variation of the maxima (dashed line) and the averages (continuous line) of collision probabilities with respect to packet generation rates in uplink and downlink. In particular, we used fixed values of the mean generation time in downlink ($1/\lambda_{down} = 1, 2, 3, 4$ hours) and drew the variation of collision probability according to λ_{up}. This figure shows that the collision probabilities do not undergo large variations as the traffic generation rate changes. For instance, we found that the mean of the collision probability when $1/\lambda_{down} = 1$ hour is 0.22% at $1/\lambda_{up} = 4$ hour and 0.52% at $1/\lambda_{up} = 30$ minutes.

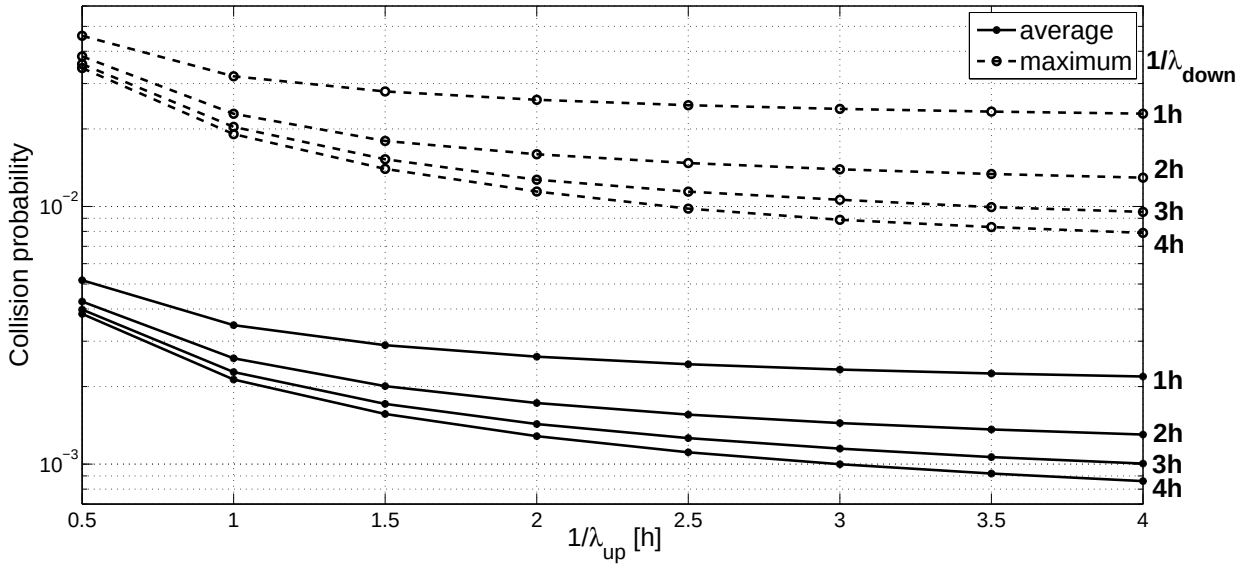


Figure 5.4 Analysis of the collision probability according to λ_{up} with fixed values of λ_{down}.

5.5.2 Impact of FHSS

Numerical results on the collision probability in different traffic scenarios (reported in Table 5.1) were presented in Malandra and Sansò (2015). In that paper it was shown that, for high traffic scenarios, the collision probabilities reached values close to one.

In order to highlight the impact of FHSS protocol on the performance analysis results, Figure 5.5 reports a comparison of the collision probabilities found with FHSS (in gray) against those presented in Malandra and Sansò (2015), without FHSS (in black). For the sake of clarity in the comparison, traffic scenarios IDs are used in this figure. A reduction of collision

Table 5.1 Traffic scenario ID according to λ_{up} and λ_{down}, taken from Malandra and Sansò (2015).

$\frac{1}{\lambda_{down}}$ [h]	$\frac{1}{\lambda_{up}}$ [h]							
	0.5	**1**	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**
1	1	5	9	13	17	21	25	29
2	2	6	10	14	18	22	26	30
3	3	7	11	15	19	23	27	31
4	4	8	12	16	20	24	28	32

probability greater than an order of magnitude as found in all the scenarios; therefore we can safely state that FHSS has a key impact on the performance of large scale RF-mesh system.

5.5.3 Delay

In Subsection 5.4.3, we explained how the delay is calculated from the probability of collision in PeRF-mesh.

In Figure 5.6, the variation of the delay in uplink is reported according to different values of traffic generation rate. In particular, we fixed the downlink mean generation time to 1 hour and let the $1/\lambda_{up}$ vary from 5 minutes to 4 hours.

On the left side of the curve, for lower mean generation times (and consequently higher traffic generation rates) there is a slight variation in the delay : we observe a mean value of 12.27 seconds with $\lambda_{up} = 5$ minutes and of 11.84 s with $\lambda_{up} = 10$ minutes, which results in a variation of -3.5%. On the other hand, the last two mean values of the delay are 11.5 and 11.49 seconds, with a variation of only -0.087%.

The flattening of the curve depends on the lower impact of collision probability on the delay as the traffic decreases. As the mean packet generation time increases, the value of the probability of collision is so low that it does not have an impact on the delay. In such cases, the delay of a packet from node i to node j, as showed in (5.7), tends to $\tau|\rho_{ij}|$ where $|\rho_{ij}|$ is the number of hops from node i to node j.

5.6 Conclusions and future steps

In this work we presented PeRF-mesh, an analytic tool to study the performance of large-scale RF-mesh systems with FHSS. To our knowledge, this is the first analytic tool to take into account the interaction of FHSS and ALOHA MAC access in a performance analysis study.

Performance analysis is key to assess the feasibility of real smart grid applications and it has some advantages with respect to stochastic simulations and real-field measurements.

PeRF-mesh allows thorough analysis of large-scale RF-mesh systems with a short computational time. Analysis of collision probability, delay and critical nodes can also allow to identify possible bottlenecks of the system in the design phase, resulting in high economical and resources savings.

The impact of the FHSS protocol was highlighted by a comparison of the numerical results obtained with PeRF-mesh against those obtained with a model without FHSS and available

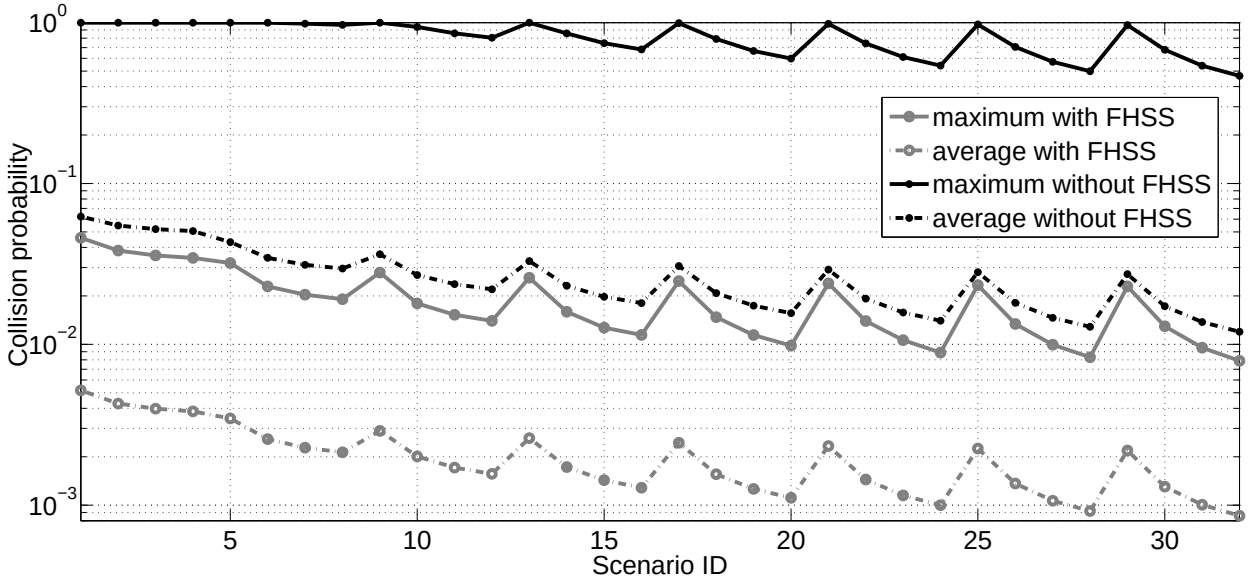


Figure 5.5 Comparison of collision probabilities with and without FHSS.

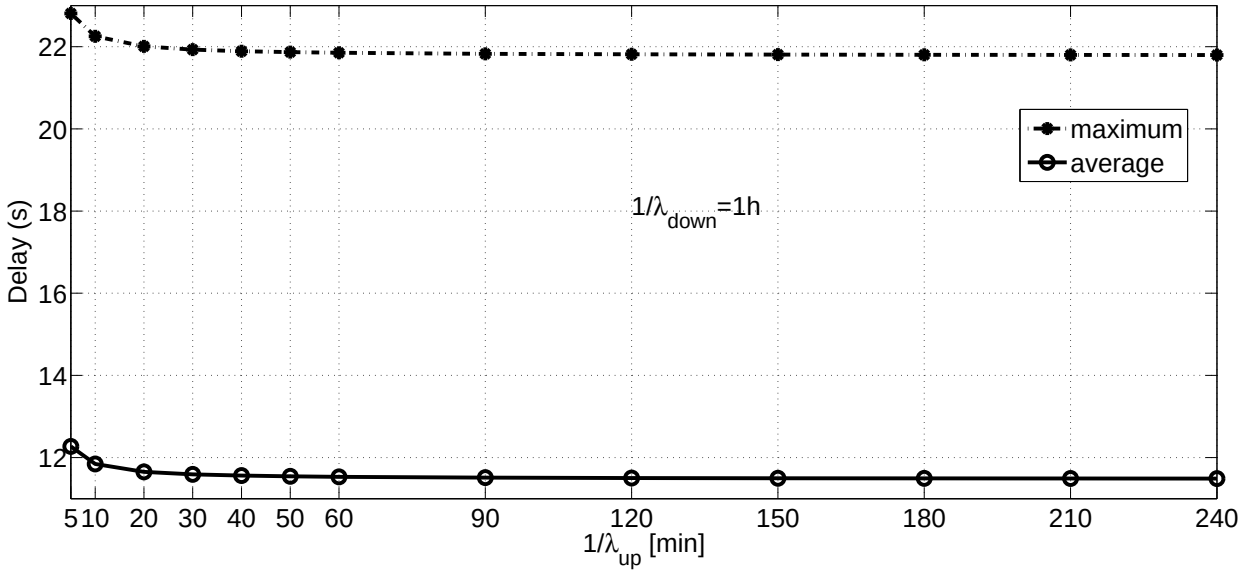


Figure 5.6 Variation of the delay according to λ_{up} with a fixed value of $\lambda_{down} = 1h$.

in Malandra and Sansò (2015). A substantial improvement was observed, in terms of a reduction in the collision probability and the consequent decrease in the delay.

One of the future steps consists in the refinement of the analytic model, investigating the possible use of a Markov modulated system; in spite of increasing the model's complexity, this can represent additional features of real RF-mesh systems, not considered so far (e.g. probability of re-transmission). Other paths to explore are the integration of more complex propagation models and of more dynamic routing protocols. Finally, a combination of optimization and performance analysis is in the agenda : we are currently conceiving a model for the optimal placement of routers and data collectors.

Acknowledgment

This work was partially funded by an ECO Energy Innovation Initiative grant from Natural Resources Canada.

CHAPTER 6 ARTICLE 3 : A SIMULATION FRAMEWORK FOR NETWORK PERFORMANCE EVALUATION OF LARGE-SCALE RF-MESH AMIs

F. Malandra et B. Sansò, (2016), "A Simulation Framework for Network Performance Evaluation of Large-Scale RF-mesh AMIs", Simulation Modelling Practice and Theory, submitted to.

Abstract

AMIs lay the foundation of a wide variety of smart-grid applications. Despite their low capacity, RF-mesh systems are an increasingly popular option for AMI implementation. In this paper, we present a simulation framework for realistic RF-mesh AMIs with the objective of analyzing the performance of this kind of system, and of defining the feasibility of possible smart-grid applications. The framework contains a simulation tool, implemented in Java and Python, that permits the simulation of large-scale AMIs, while taking into account important technical details in a reasonable amount of time. Numerical results obtained in a real case scenario are also proposed and discussed.

6.1 Introduction

RF-mesh Advanced Metering Infrastructures (AMIs) are increasingly popular and can be used as a transmission support for different types of smart-grid applications. However, an accurate simulation of RF-mesh AMIs and a proper evaluation of their performance are fundamental to assess the behaviour of the system when all those applications are functional.

Network simulators are usually designed to recreate the performance behaviour of a system taking into account network layer protocols and conditions. Therefore, when working with RF-systems, cross-layer modules have to be integrated into commercial and/or open source simulation tools, in order to be able to study the network layer performance while considering the lower layer functionalities and phenomena. Reviewing the literature on RF-mesh AMI simulators, we found that most works relied on OPNET Bian et al. (2014); Chen et al. (2013, 2014); Islam et al. (2014); Panchadcharam et al. (2012); Patel et al. (2011), while others employed NS-2 or NS-3 Aalamifar et al. (2014); Rice and AlMajali (2014); Yigit et al. (2013), Contiki OS and its network simulation module Cooja Ancillotti et al. (2013); Elyengui

et al. (2015); Gormus et al. (2014), CASTALIA Iyer et al. (2011), and OMNET++ Iyer et al. (2013); Lichtensteiger et al. (2010).

OPNET is a commercial network simulator with a user-friendly Graphical User Interface (GUI) and a wide variety of modules to simulate different wired and wireless technologies and protocols. The authors who employed OPNET carefully took into consideration some aspects of AMI implementations, such as routing, propagating condition, QoS requirements; however, the network sizes considered in the published studies were too small to represent the large-scale phenomena of real life RF-mesh AMIs Bian et al. (2014); Chen et al. (2013, 2014); Islam et al. (2014); Panchadcharam et al. (2012); Patel et al. (2011).

NS-2 and NS-3 are two very popular open-source discrete-event network simulators that allow for an accurate representation of the physical channel with complex propagation models (i.e., Nakagami-Rayleigh in Yigit et al. (2013) and HATA COST231 in Aalamifar et al. (2014)). Again, only small-size networks (of around 60 nodes) were considered in the two contributions and the time horizon employed was too short (e.g., 20 s in Aalamifar et al. (2014)) to provide general considerations on AMI performance. The authors of Rice and AlMajali (2014), on the other hand, used 350-node instances but their analysis was focused on the performance of an AMI under a cyber-attack; therefore, they did not deal with general application performance issues.

Contiki is an open-source OS, mainly used to connect small low-cost, low-power micro-controllers to the Internet. It supports a wide range of wireless standards and communication protocols. It also provides Cooja, a network simulator particularly used in the performance analysis of AMI systems. Unfortunately, no module dedicated to RF-mesh technology was found in the literature Ancillotti et al. (2013); Elyengui et al. (2015); Gormus et al. (2014).

CASTALIA, an open-source discrete-event-driven network simulator, was used in Iyer et al. (2011) to analyze the AMI performance. The instance size was again in the order of few hundred nodes, and the geographical (e.g., location of nodes) and topological (e.g., number of neighbors) aspects of the problem were not considered. Moreover, the performance analysis only encompassed the hop count and an end-to-end delay histogram.

The free discrete-event-driven network simulation software OMNET++ was used in Iyer et al. (2013), where different routing solutions were compared. Despite the large-scale nature of the instances, in Iyer et al. (2013) GIS data were not used and the positions of smart meters were randomly distributed in a given area. An inspiring work using OMNET++ is Lichtensteiger et al. (2010), where GIS data were used to build consistent large-scale instances; the authors, however, mainly focused on hop count statistics, providing little detail on some important

performance indexes, such as latency or the amount of time spent in transmission by smart meters.

Despite the fact that existing simulation platforms are extensively employed in general networking, some issues and gaps were observed in the RF-AMI performance literature. First, there is a lack of modules dedicated to the RF-mesh technology in affordable off-the-shelf software. Second, some important modelling details such as FHSS (which will be explained in Section 6.2.2) are not implemented. Third, there is an issue of low scalability, as most of the literature papers simulate wireless systems up to few hundreds of smart meters, a consistently lower scale compared to currently implemented RF-mesh AMIs. Moreover, many articles do not focus on the geographical and topological aspects of the problem, which are important to better analyze the peculiarities of RF-mesh AMIs. Finally, there is a lack of publicly available information, which makes it difficult to compare and reproduce different situations and results : some authors do not share details of their simulations, others use software that is not open-source.

As a conclusion, the study of the literature on AMI simulators shows that modifying general purpose networking simulators leads to some important shortcomings that do not allow to capture the performance of large-scale systems. Therefore, we decided to part from the aforementioned studies and avoid implementing yet another module of an existing network simulator. Rather, we propose a new framework that captures the most important performance features while allowing the study of large-scale AMI-systems.

To fill the gap in currently available solutions, we propose a simulation framework with the aim of achieving a detailed performance analysis of large-scale RF-mesh AMI systems. Our methodology permits to calculate the probability of collision of packets and the end-to-end delays in different geographical areas, analyzing possible bottlenecks of the system. Wireless interference and its negative effects on performance are carefully accounted for, as well as other performance indexes, such as the activity time, that can be interesting to limit the public fears on RF exposure. The framework is highly scalable and allows the study of large-scale geographical regions with thousands of smart meters, running different types of applications.

The remainder of this document is structured as follows : the main features of a RF-mesh system are presented in Section 6.2; the simulation framework is introduced in Section 6.3, and numerical results obtained with the developed simulation tool are presented in Section 6.4; Section 6.5 contains the conclusions and some discussion.

6.2 Description of the RF-mesh system

6.2.1 Architecture

Figure 6.1 shows that RF-mesh AMIs have a layered architecture divided as follows :

- HAN. This includes electric vehicles, space heaters, water heaters and all the smart appliances within a home and it is characterized by short range communications. All the involved nodes communicate with the smart meter, which acts as a gateway to the upper layers. The main adopted communication technologies are ZigBee and Bluetooth, but also WiFi and Power Line Communications (PLC) are being increasingly used.
- NAN. This is a wireless mesh network in which all the smart meters are connected to a data collector. Wireless routers are also installed in order to increase the connectivity and extend the area coverage. The adopted technology is RF-mesh, in which the ISM band of 900 MHz is used to transmit data on a mesh topology.
- WAN. This is the IP-based backhaul of the communication system : data collectors communicate with the power utility MDMS by means of satellite and cellular transmissions.

The first and the third layers of the architecture (i.e., HAN and WAN) use well-known technologies (e.g., ZigBee or cellular) whose performance has already been extensively investigated in the literature. On the other hand, RF-mesh-system performance has not been as thoroughly studied, and is the topic of this study. The NAN layer of the system under study is composed of :

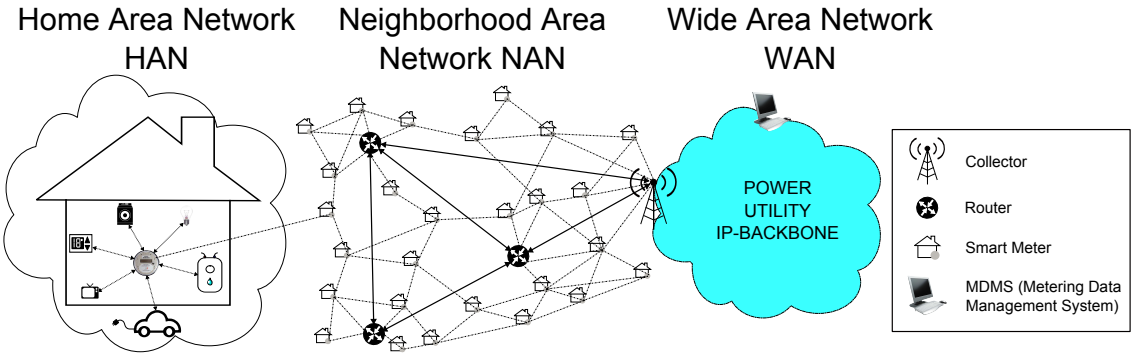


Figure 6.1 Architecture of the RF-mesh AMI.

- *Data collectors*, that are gateways between the NAN and the WAN. They have enhanced radio capacities (e.g., bandwidth, covering ray) with respect to the other nodes. They produce packets directed to the smart meters at a rate that depends on the implemented smart-grid applications.
- *Routers* are devices used as relays between smart meters or between a meter and a data collector. They do not produce their own packets, but are equipped with stronger radio transmitters than those of the smart meters.
- *Smart meters* represent the gateways between the HAN and the WAN; they produce packets towards the data collector at a rate that depends on the implemented smart-grid applications.

6.2.2 Main features

The main features of the NAN of a RF-mesh AMI to be taken into account in the simulation framework can be summarized as follows.

Large number of nodes

Smart meters are the most numerous devices in the topology : their number ranges from roughly a thousand per neighborhood in a rural environment up to tens of thousands in densely populated urban environments. The number of routers varies according to the considered scenario : in urban instances, very few routers are required since the connectivity of the network is already high, while in rural environments a higher number might be necessary in order to have a fully connected topology, since nodes are otherwise far from each other. The high number of nodes also affects the size of the simulations, increasing the computational burden as the network size increases.

Low throughput

Several factors (e.g., the interference, the use of public frequency bandwidth, and low quality devices) make the nominal rates of the wireless links low, i.e. 19.2 kbps for links between routers and data collectors, and 9.6 kbps for other links (data from Lichtensteiger et al. (2010)). The low throughput hinders the use of RF-mesh AMI for applications that require higher data rates (e.g., video surveillance) but the system under study remains a good candidate for many other applications (e.g., meter reading, load management), which do not need large data transfers.

Black-box nature of the system and undisclosed features

AMI systems are often sold to power utilities as real black-boxes, and very technical details are only provided under strict non-disclosure agreements. Moreover, once installed, smart meter data is considered very sensitive. As a result, some of the characteristics of RF-mesh AMIs are not publicly available, despite their importance for the analysis of the network performance. In the process of designing our framework, we did some reverse engineering and carefully scanned any publicly available piece of information to make the simulator more realistic, while developing the modules as flexible as possible to encompass many different types of undisclosed features.

MAC layer

Two packets that are simultaneously transmitted on the same wireless channel create a *collision*. Wireless systems react in different ways after a collision, depending on the implemented MAC layer protocol. In our simulation framework, we considered a smart-meter communication system using a MAC layer based on time-slotted ALOHA. According to this protocol, the time is divided into time slots and a node is allowed to transmit at the beginning of each time slot : therefore, when a node has a packet ready, it waits until the beginning of the following time slot to transmit it. If the correct reception of a given packet i by node A is prevented by a concurrent transmission on the same time slot, node A is *backlogged* and will attempt a retransmission of packet i in one of the following time slots with probability p_r. Despite reducing collision occurrences, the introduction of the retransmission probability degrades the overall performance because it increases the system delay.

Wireless interference issues

As was previously mentioned, AMI RF transmissions take place in a free unlicensed band, where a large number of devices can simultaneously transmit, leading to severe interference issues.

The presence of thousands of potential simultaneous transmitters in a given area requires the adoption of dedicated protocols to mitigate the effect of interference. Among the available protocols, FHSS is one of the most frequently adopted in actual RF-mesh AMIs implementations. The protocol, as shown in Figure 6.2, consists in subdividing the frequency spectrum into θ sub-frequencies. Signals are then transmitted using just one frequency channel at each time slot. The frequency channel used for transmission varies at each time slot, according to a predetermined sequence that is known and used by all nodes. Two devices interfere with

each other if and only if they transmit on the same frequency channel at the same time slot. In order to avoid this, the sequence is shifted over time in each node, in order to reduce the number of neighbors with synchronized sequences. More details on the FHSS implementation can be found in Section 6.3.3.

FHSS is particularly efficient against the interference coming from other sources (e.g., cordless phones, remote controllers) : the power of the interfering signals is spread over the entire available spectrum and their effects on the receiver side are considerably limited. Devices in the same RF-mesh system are considered the only possible sources of interference.

Network layer

An active role in the dynamics of a mesh system is played by the routing protocol, that can greatly impact the network performance. To the best of our knowledge, there is no one prevailing routing mechanism and details about the routing in place in implemented systems are scant. Geographical routing seems to be one of the most popular mechanism in RF-mesh AMIs, but it requires each node to be equipped with a GPS antenna and to know the coordinates of all the other nodes. The so-called *Layer-based* routing is also used in existing systems. In this routing protocol, the word *layer* is used to identify the hierarchical division of the nodes. Every node is assigned a layer number (i.e., collectors are indicated with 0, their neighbors with 1 and so on). The downlink path is decided by the collector based on the information collected in the layer formation phase. On the other hand, each smart meter transmits packets in the uplink direction using one of its neighbors in the upper layer. The advantage of this mechanism is that it is very simple on the smart meter side. Additional details about the routing implementation are provided in Section 6.3.2.

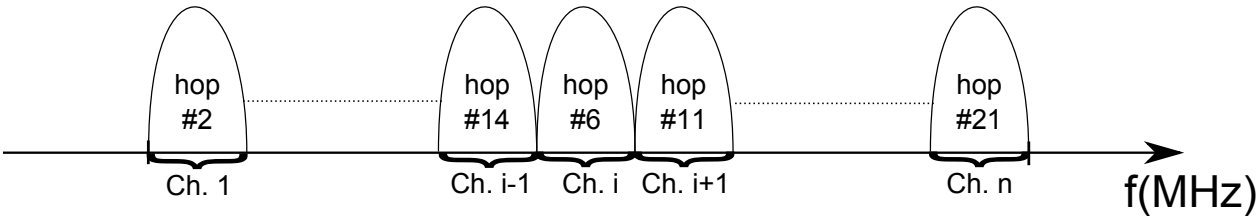


Figure 6.2 Example of operation of FHSS protocol.

6.3 Simulation framework

6.3.1 General structure

In Figure 6.3, the general structure of the simulator is reported. Three phases are identified : (i) the initialization phase (described in Section 6.3.2), (ii) the simulation phase (described in Section 6.3.3), and (iii) the results analysis phase (described in Section 6.4).

6.3.2 Initialization phase

The first phase of the simulation framework, schematized in Figure 6.3, is the *initialization*, whose main objective is to prepare the instance to be simulated. The user is given the possibility to choose the instance that best reflects his/her needs. The flexibility of this phase is key to enable a broad set of performance analyses. Figure 6.4 shows that the initialization phase encompasses the nodes placement (described in Section 6.3.2), the link definition (described in Section 6.3.2), and the routing (described in Section 6.3.2).

Let us define some mathematical notation that will be used in the rest of this section :

V	Set of nodes in the topology
M	Set of smart meter nodes
R	Set of routers
C	Set of data collectors
E	Set of links in the topology
$d(i,j)$	Euclidean distance between node $i \in V$ and node $j \in V$
R_m	Covering ray of each smart meter
R_r	Covering ray of each router and collector
$\nu(i)$	Set of neighbors of node i

Nodes placement

As was already mentioned, three types of devices are used in RF-Mesh AMIs : the smart meters M, the routers R, and the data collectors C. It is important to remark that every device $i \in V$ has a unique functionality and cannot have two different roles nor change its role in the course of a simulation : $M \cup R \cup C = V, M \cap R = M \cap C = C \cap R = \varnothing$. Every node is characterized by : 1) its role, 2) a unique integer ID, and 3) its GPS coordinates. The ID notation follows the convention that the first node is a collector with ID equal to 0. The

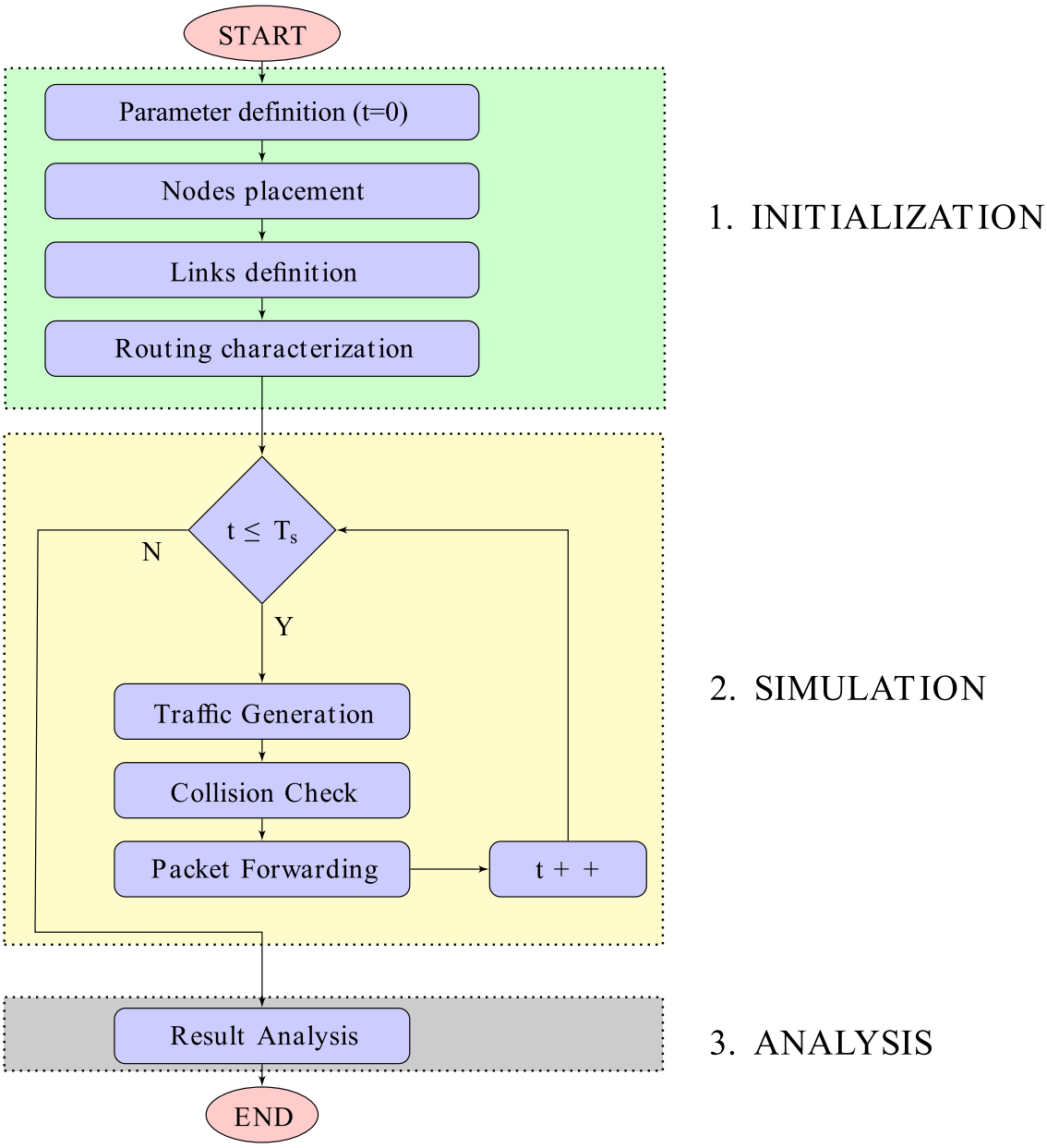


Figure 6.3 Simplified architecture of the simulator.

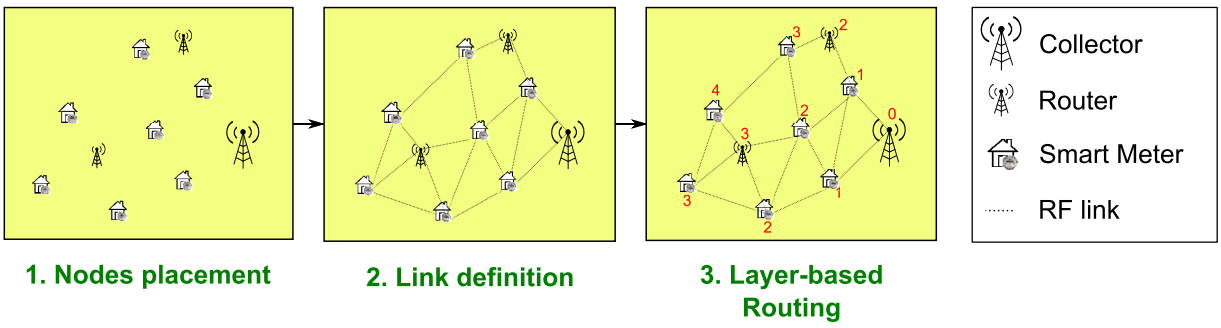


Figure 6.4 A scheme of the initialization phase with a simple topology with 10 nodes.

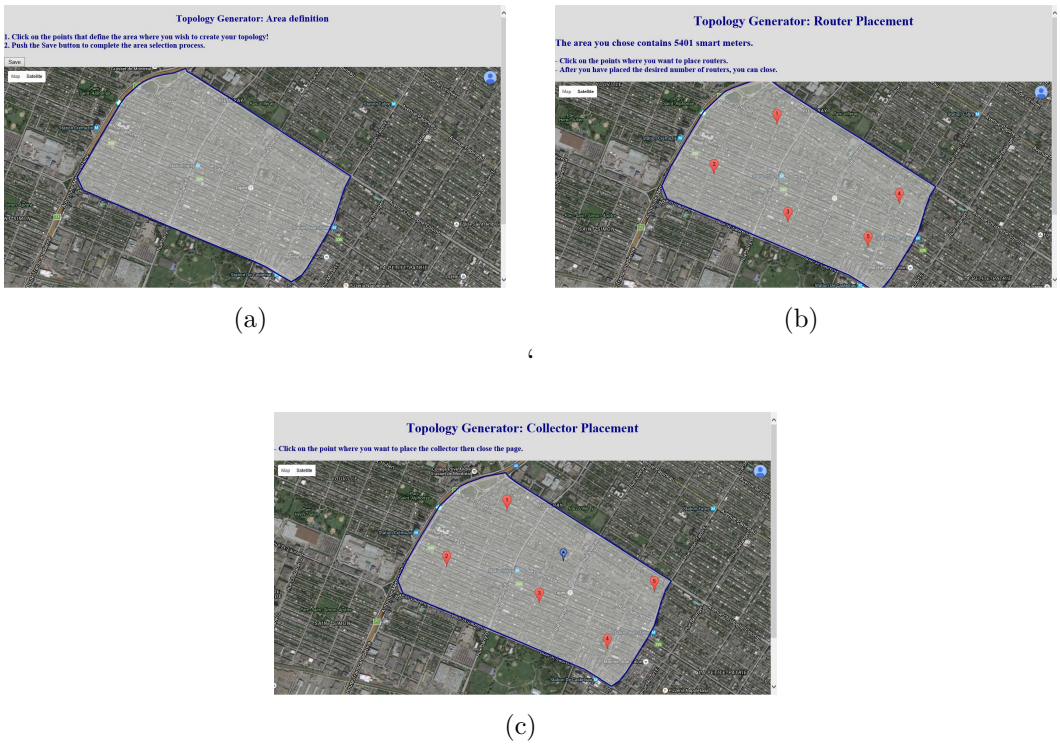
other IDs are incrementally assigned, starting with any other data collectors, then following with the routers, and finally with the smart meters. At least one data collector is required but, as shown in the rest of the paper, the number of data collectors might be greater than 1.

Nodes position can be either manually selected in a map by means of a geographical tool we implemented, or directly provided in a text file. The manual selection process does not apply to smart meters, for their number is too large (i.e., on the order of thousands); it is rather more convenient to retrieve their position by one of the various portals providing open GIS data (data for the Montreal island, used in our numerical results, can be retrieved from Municipality of Montréal (2014)). On the other hand, the position of routers and collectors can be not only selected on a map, but also optimally determined by integrating a facility location model into the current framework, which we are currently studying.

In Figure 8.1, an example of the topology creation procedure is illustrated. The user can draw a polygon to define the analyzed area (see Figure 8.1(a)), then we implemented a script that allows to identify the smart meters located within the selected area; finally, the user is required to select on the map the position of routers (see Figure 8.1(b)) and data collectors (see Figure 8.1(c)).

Link definition

The topology is completed with the definition of the set E of links. In order to keep the complexity low, the definition of the links is based on the euclidean distance between nodes. Two different covering rays were defined : R_m for the smart meters, and R_r for the routers and collectors. Numerical values of R_m and R_r vary according to the scenario considered : the covering rays are typically shorter in urban scenario than in rural environments, due to the harsh propagating condition (i.e., many many obstacles exist in the path) and the



(a) (b)

(c)

Figure 6.5 Illustration of the topology creation with the definition of a new area (Figure 8.1(a)), the choice of routers (Figure 8.1(b)), and collectors (Figure 8.1(c)).

concurrent transmissions of other wireless users in the same unreserved bandwidth. Moreover, the covering ray of the routers and collectors is considerably higher than the covering ray of smart meters because of the use of better antennas and also because routers and collectors can exploit better propagating conditions (they are usually installed on the top of buildings).

E is the set of links (i, j) characterized as follows :

$$E = \{(i,j) : (i,j \in V) \wedge (d(i,j) \leq R_m \wedge i,j \in M) \vee (d(i,j) \leq R_r, (i \in R \cup C) \vee (j \in R \cup C))\} \tag{6.1}$$

where $d(i, j)$ is the euclidean distance between node i and j, calculated using the GPS coordinates of the nodes according to the Cosine-Haversine formula Robusto (1957). In other words, there is a link between i and j if they both are smart meters and their distance is lower than R_m, or if at least one of them is a router or a collector and their mutual distance is lower than R_r.

Routing

In RF-mesh systems, packets flow in just two directions : uplink (from a smart meter to a data collector) and downlink (from a collector to a smart meter). The communication between two smart meters is not possible for confidentiality reasons : messages are encrypted with a key that is shared between the data collector and each smart meter. The routing mechanism is the so-called *layer-based routing*, similar to the one presented in Chen et al. (2012). An initial *layer-formation* process is required to assign a *layer index* to each node. The *layer-based* routing relies on the *layer-updates*, small packets that are used in order to maintain the routes (from smart meters to data collectors and vice versa) up-to-date and to react to possible changes in the topology (e.g., link losses, node failures). For instance, in case a given node i with layer k attempts a transmission to a $(k-1)$-layer node, which becomes unavailable, i will choose one of its neighbors with the lowest layer number, and a path to the collector can be promptly re-established. In case no $(k-1)$-layer neighbor is active, node i chooses one of the neighbors with a higher layer index.

For the sake of illustration, in Figure 6.4, a scheme of the *initialization* phase is reported for a simple topology with 10 nodes. As shown in the third subfigure, a layer index is assigned to each node. The path in the downlink direction is determined by the collector, based on the updated information received in the *layer updates*. In the uplink direction, the path is not known in advance by the source of the communication : a given smart meter with layer index k transmits the uplink packet to one of its $(k-1)-$layer neighbors. This routing mechanism is a good candidate for AMI systems because it does not require a high computational burden at the smart meter side. Nevertheless, it is dynamic, robust and capable of reacting to topology changes.

6.3.3 Simulation phase

A scheme of the logical structure of the nodes is reported in Figure 6.6 : each node is equipped with 1) two logical antennas, 2) a FHSS module, 3) two separate buffers (one for incoming packets and one for packets waiting to be transmitted), 4) a packet analyzer, 5) a routing mechanism. However, the elements in the grey area (i.e., destination and traffic generator) are not implemented in routers, but only in smart meters and collectors, which can be sinks or sources of traffic. The two antennas (one for receiving and one for transmitting) are modeled as two logical entities associated to the same physical antenna each node is equipped with. The presence of two logical antennas in the simulator is needed to allow for the possibility to work with one frequency in reception and another in transmission.

Let $F_r(j,t)$ and $F_t(j,t)$ be the frequency channels of the two antennas (respectively, receiver and transmitter) of node j at time t. $F_r(j,t)$ is determined according to the FHSS sequence, known to all the nodes. A given node i, at the beginning of a time slot, can be either in *idle* state or in *transmission* state. When in idle state, it is waiting to receive a packet and its antenna is tuned to the frequency $F_r(i,t)$. On the other hand, when in transmitting state, the transmitting frequency of node i at time t ($F_t(i,t)$) depends on the destination of the transmission : if node i wants to transmit a packet to node j at time t, it must tune its transmitting frequency to the receiving frequency of node j at time t, namely $F_t(i,t) = F_r(j,t)$. This mechanism is implemented to reduce the chances of collision : a collision in the transmission from node i to node j happens when another neighbor of node j is transmitting on the same channel $F_r(j,t)$.

The simulation phase, as shown in Fig. 6.3, consists in the iteration, over the simulation horizon T_s, of three sub-phases (i.e., traffic generation, collision check, and packet forwarding) that are described in detail in the next sections. The simulation phase is *time-driven*, meaning that T_s is subdivided into time slots of duration τ, and a variable is used to record the current time, which is incremented as all the aforementioned phases are performed. An event-driven approach is usually preferred when the number of events is considerably lower than the number of time steps in the simulation. However, the large number of nodes and possible events in the system under study (e.g., reception/transmission of a packet, collisions, buffer updates) made us prefer a time-based approach.

The *traffic generator* module

At the beginning of each time slot, all the nodes that are equipped with a *traffic generator* can produce a packet according to a random distribution that depends on the type of the implemented applications.

Let us denote by $\Omega = \{\alpha, \beta, \gamma\}$ the set of three different traffic types in the communication system under study. In this case, we have considered meter-reading, demand, and broadcast traffic.

- α (*meter-reading*) consists in the transmission of a packet from a smart meter to its associated collector. This event is Poisson-distributed with parameter λ_u.
- β (*demand*) consists in the random transmission of packets from a data collector to one of its controlled meters. The generation of a packet is also Poisson-distributed (mean parameter λ_d).

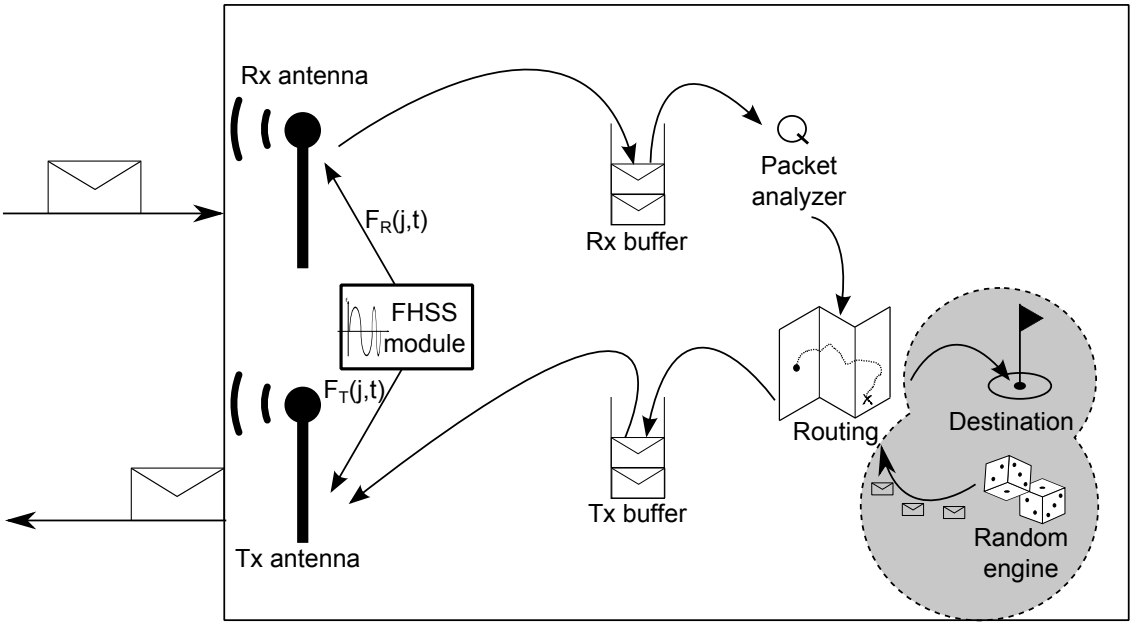


Figure 6.6 Logical structure of the different types of nodes in the simulator.

γ (*broadcast*) consists, in general, in the transmission of the same packet from a data collector to all smart meters. In practice, broadcast transmissions, not possible because of privacy requirements, are replaced by unicast packets; nonetheless, in the rest of this paper we will refer to γ as a *broadcast* transmission.

Note that α flows in the uplink direction, while β and γ take place on the downlink. In Figure 6.7 we reported a screenshot of the mask to set up the initial parameters of the simulation. The three traffic sources (i.e., α, β, and γ) can be selected by means of the corresponding check-boxes, which were circled in the illustration.

The analysis of different kind of applications in the same communication infrastructure can be handful from the point of view of a power utility. Smart grid is constantly enlarging its application domain and the possibility of using only one communication system for a broad set of applications is particularly sought by power utilities.

The *collision check* module

When a given node i is attempting a transmission to one of its neighbors $j \in \nu(i)$, possible interferers of the transmission $i \to j$ are all the nodes $k \in \nu(j) : k \neq i, j$ who transmit on the same frequency channel as i. In order to identify the collisions, the dedicated *collision-check* module was implemented. At each time slot, a special check is made to see if any

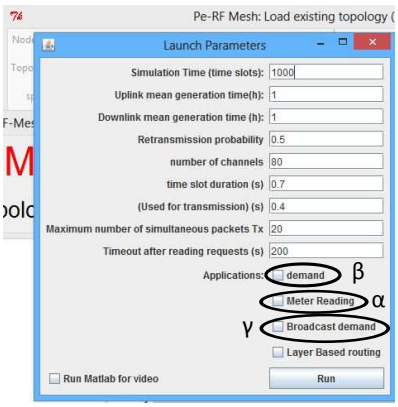


Figure 6.7 Mask used to set up initial parameters to launch a simulation. The different traffic generation parameters are circled and labeled with the corresponding letter, defined in Section 6.3.3.

of the transmitting nodes is interfering with each other. Each transmission is identified by a boolean *collision-flag* variable, whose default value is *false*. When a packet collides, its *collision-flag* variable is set to *true* and the packet is scheduled for retransmission at one of the following time slots.

The *Packet-forwarding* module

At the end of each time slot, after the *traffic generator* module has decided which nodes are transmitting and after the *collision-check* module has determined the presence of collisions, it is necessary to update the Tx and Rx buffers. For this, we implemented the *Packet-forwarding* module. For each transmission from node i to node j, this module checks if a collision has been reported. Then, in case of collision, the packet stays in the Tx buffer of node i and a new transmission will be retried in one of the following time slots with probability p_r. In case no collision was observed, the packet is deleted from the Tx buffer of node i and forwarded to the Rx buffer of node j. Then the *packet-analyzer* (see Figure 6.6) of node j verifies the destination of the packet : if node j is the destination, the packet is considered successfully received and packet statistics are updated; if node j is not the destination, the packet will be passed to the *routing* module, in order to determine the next hop, and written to the Tx buffer of node j.

6.3.4 Randomness in the simulator

Random number generation is a key element in computer-aided simulations and is usually performed by means of the so-called *random engines*, objects that can generate numbers according to different types of distributions (e.g., uniform, beta, exponential). Those engines, regardless of the operating system and of the used programming language, adopts computational algorithms to generate pseudo-random sequences, which are based on an input parameter, called *seed*. The use of random engines and initial seeds guarantees the possibility of re-running simulations, and enables analyses on the truthfulness of the simulation results. Figure 6.8 illustrates how the random engines of the nodes are initialized. As shown, a central random engine is used to generate the seeds for all the random engines in the simulator. Each random engine generates traffic according to the mean packet generation rate of smart meters and data collectors (i.e., respectively λ_u and λ_d).

6.4 Numerical results

To illustrate the interest and versatility of the proposed framework, we present in this section a set of results that were obtained by analyzing collision probability, delay and smart-meters activity time in a realistic setting built from publicly available data.

The scenario used to produce this set of numerical results is based on Villeray, a neighborhood within the city of Montreal, which was one of the three areas chosen for a pilot project of smart meter installation in Québec Hydro-Québec (2012). The starting topology was created with the topology generator described in Section 6.3.2. In particular, the position of routers and collectors were extracted from Hydro-Québec (2012) and stored into a text file. The GPS coordinates of smart meters, not provided in the previous document, were extracted in an automated way by a dataset [1] containing all the residential addresses in the city of Montreal and stored in a text file as well. The text files with the GPS coordinates of the nodes are used as input of the topology generator. The topology is composed of 2 data collectors, 16 routers, and 6033 smart meters over an area of 2.7 km^2. In our simulations, we created 20 traffic scenarios with $1/\lambda_u = 1, 0.5, 0.25, 0.125$ h and $1/\lambda_d = 4, 3, 2, 1, 0.5$ h and no broadcast transmission. The effect of broadcast traffic on the performance was studied in 20 additional traffic scenarios with the same values of $1/\lambda_u$ and $1/\lambda_d$, but adding one broadcast transmission per day at a predetermined time slot. Relevant parameters that affect the simulation results are τ (i.e., the time slot duration), C_r and C_m (i.e., respectively the capacity of links between routers/collectors, and the capacity of all other links), L (i.e., the packet size in Bytes), and

[1] Provided by *Portail Données Ouvertes Montreal, available at* `http://donnees.ville.montreal.qc.ca/dataset/adresses-ponctuelles`

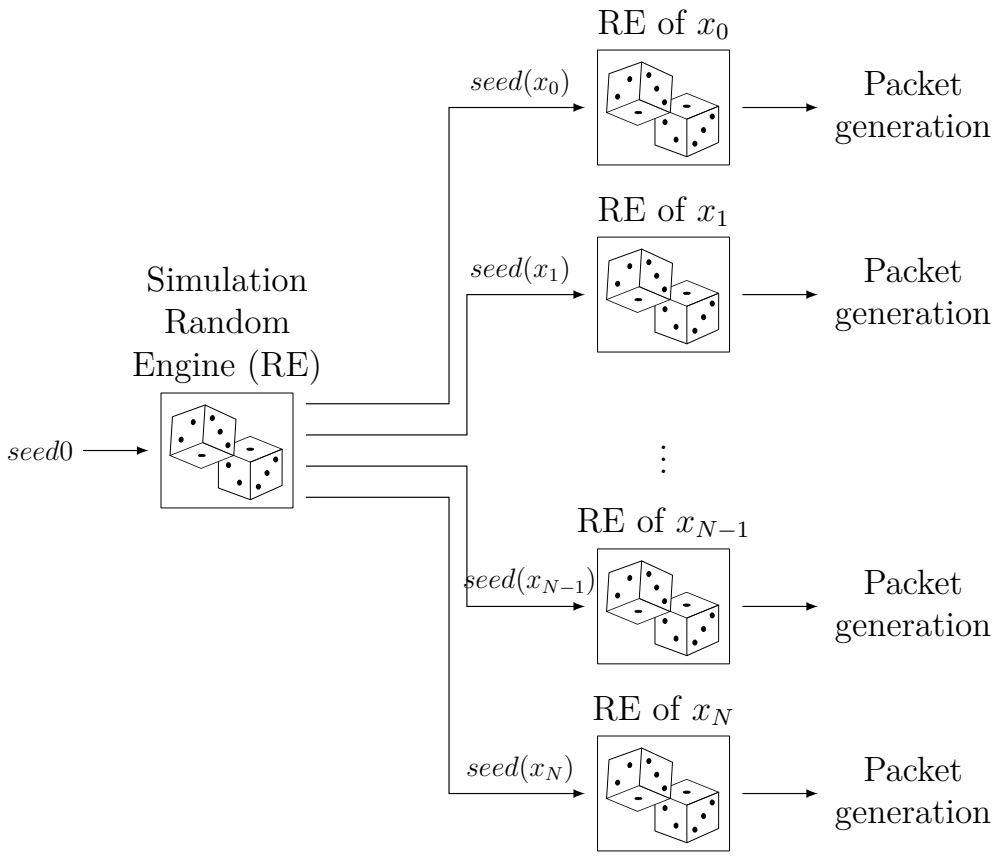


Figure 6.8 Illustration of the seed initialization and of the random packet generation in the simulator.

Θ (i.e., the buffer size expressed in number of packets). The average number of packets that can be simultaneously transmitted on the links between routers/collectors can be found as $\phi_r = C_r\tau/L$, and on the other links it is $\phi_m = C_m\tau/L$. In the case under study, given that $C_m = 9.6$ kbps, $C_r = 19.2$ kbps, $\tau = 0.7$ s and $L = 100$ Bytes $= 800$ bits, we have $\phi_r = 16.8$ packets and $\phi_m = 8.4$ packets.

Additional notation, used in the rest of this section, is reported below :

$N_t(i)$	total number of transmissions of node $i \in V$
N_t	overall number of transmissions
$N_c(i)$	total number of collisions experienced by node $i \in V$
N_c	total number of collisions
$\pi(i)$	collision probability at node i
π	overall collision probability

6.4.1 Confidence interval analysis and results reliability

For every traffic scenario, we performed 50 simulations (i.e., we used 50 different initial seeds $seed_0$) in a 7-day time horizon. The use of multiple initial seeds permits to calculate confidence intervals for the parameters that are presented in this section. If we denote by $k = 1..50$ the set of initial seeds, and by $X^{(k)}$ the observation of a random variable X when the initial seed is equal to k, we can compute :

$$m = E[X^{(k)}] = 1/50 \sum_{k=1}^{50} X^{(k)}, \qquad s^2 = var(X^{(k)}) = 1/50 \sum_{k=1}^{50} (X^{(k)} - m)^2 \tag{6.2}$$

and use them to compute the confidence interval at the 95% confidence level as $E[X^{(k)}] \pm 1.97s$. In this case, 95% represents the reliability of the estimation of the real mean μ of the random variable X, which is unknown.

6.4.2 Collision probability

One of the elements that most affect the performance of a wireless network is the mutual interference of transmitting nodes. FHSS protocol mitigates the effects of wireless interference by increasing the number of communication channels and consequently reducing the chances of collisions, as shown in Malandra and Sansò (2015); Malandra and Sansó (2015). However, collisions are not completely eliminated because of the limited number of available channels and the high number of neighbors, and consequently interferers, each node has in RF-mesh AMIs.

Numerical results on the collision probability π are summarized in Table 6.1 for the 20 simulations run with simulation time $T_s = 864010$ time slots (roughly 7 days). In the Table, the 95% confidence intervals for values for π are reported according to the different values of λ_u and λ_d used in the simulations. Note that the average collisions probability ranges from 1.77% (in the scenario with $1/\lambda_u = 1$ hour and $1/\lambda_d = 4$ hours) to 9.21%(in the scenario with $1/\lambda_u = 0.125$ hour and $1/\lambda_d = 0.5$ hours) in the scenarios without broadcast transmissions. Moreover, the introduction of one broadcast transmission per day does not seem to increase the overall the collision probabilities which range from 1.92 % to 9.28 %.

One of the simulated scenarios (i.e., with one broadcast transmission per day, $1/\lambda_d = 0.5$ h, and $1/\lambda_u = 0.125$ h) has been studied separately in the rest of this section. A more accurate portrait of the dynamics of the network in the aforementioned scenario is shown in the heat-map of the collision probabilities in Figure 6.9. The two heat-maps do not considerably differ : the slightly higher collision probability in downlink can be remarked by looking at the

Table 6.1 95% confidence intervals of the average collision probability (π) according to different levels of $1/\lambda_u$ (rows) and $1/\lambda_d$ (columns).

$1/\lambda_d$ (hours)	Collision probability π (%) $1/\lambda_u$ (hours)			
	1	0.5	0.25	0.125
4.0	**1.77±2.38e-1**	3.08±4.08e-1	5.13±8.43e-1	7.50±9.76e-1
3.0	1.93±2.89e-1	3.26±3.96e-1	5.31±7.90e-1	7.75±1.12
2.0	2.22±3.77e-1	3.51±4.84e-1	5.66±7.94e-1	7.93±1.20
1.0	2.83±3.72e-1	4.22±6.02e-1	6.38±9.67e-1	8.84±1.37
0.5	3.30±4.51e-1	4.67±7.25e-1	6.78±9.81e-1	**9.21±1.39**
$1/\lambda_d$ (hours)	Collision probability π (%) - with broadcast traffic $1/\lambda_u$ (hours)			
	1	0.5	0.25	0.125
4.0	**1.92±3.09e-1**	3.12±4.74e-1	5.16±7.32e-1	7.48±9.15e-1
3.0	2.06±3.01e-1	3.26±4.06e-1	5.40±7.25e-1	7.76±1.14
2.0	2.30±3.53e-1	3.54±5.27e-1	5.57±8.25e-1	7.91±1.02
1.0	2.89±4.23e-1	4.29±6.24e-1	6.44±9.03e-1	8.86±1.40
0.5	3.29±4.86e-1	4.66±7.91e-1	6.95±1.01	**9.28±1.54**

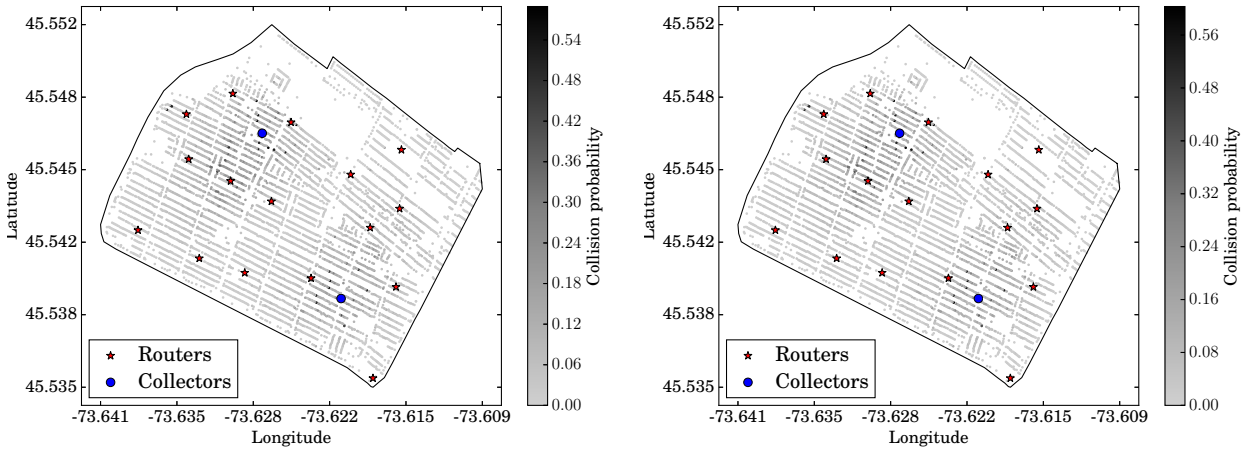
difference between the two color-bars. The perimeter of the analyzed area was also drawn in order to highlight the area analyzed in the simulation. Even though the average collision percentage of all nodes is always low, according to Table 6.1, from Figure 6.9 we can see that the situation is considerably different in the surroundings of the 2 data collectors : a high number of collisions is experienced and very high collision probabilities are observed. The reason behind the increased number of collisions is to be found in the higher concentration of packet transmissions, in both uplink and downlink directions, in that area.

6.4.3 Delay analysis

Delay is one of the most important indexes of network performance. If we denote with $T_d(k)$ the generation time slot of packet k and with $T_a(k)$ the arrival time slot at its destination, we can generally define the delay $D(k)$, encountered by packet k, throughout the network as :

$$D(k) = T_a(k) - T_d(k) + 1 \tag{6.3}$$

It is also important to separately study the average delay for the uplink and downlink traffic streams, respectively $D_\alpha(i)$ and $D_\beta(i)$. As was done for the collision probability, we



(a) without broadcast

(b) with broadcast

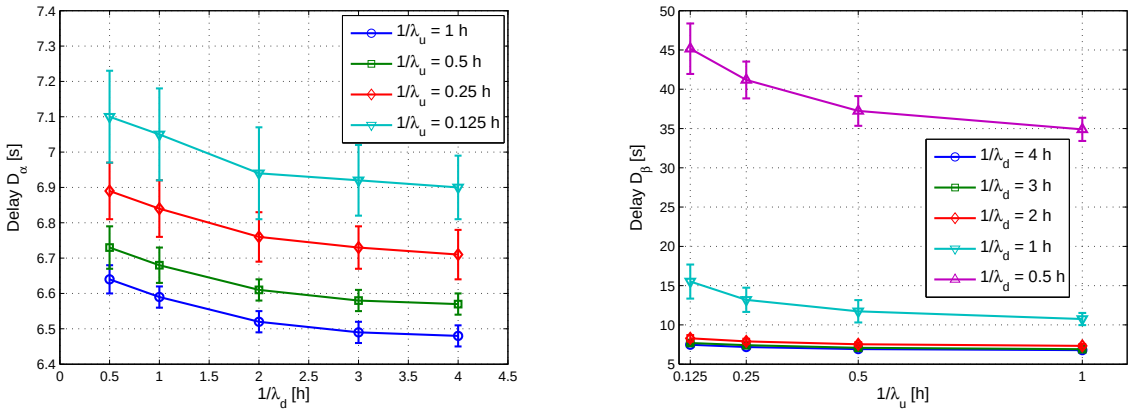
Figure 6.9 Heat-map of the collision probability in the scenario with $1/\lambda_d = 0.5$ h, and $1/\lambda_u = 0.125$ h.

carried out a 95% confidence interval analysis of the delay not considering broadcast traffic. In the upper part of Table 6.2, we reported the values of D_α and D_β without broadcast transmissions, according to the previously defined ranges of variation of $1/\lambda_d$ and $1/\lambda_u$. As one can see, the delay in uplink (D_α) slightly decreases as the packet generation rate increases, but in the downlink direction (i.e., D_β) higher delays are experienced, as also illustrated in Figure 6.10. In particular, we can notice that when the mean packet generation time in downlink goes below 1 h, there is a steep increase in the average downlink delay, with values ranging from 34.89 s to 45.16 s.

In Figure 6.11, we reported the two heat maps of the delays (uplink in Figure 6.11(a) and downlink in Figure 6.11(b)) in the scenario with one daily broadcast transmission, $1/\lambda_d = 0.5$ h and $1/\lambda_u = 0.25$ h. Note that the average delay in the downlink direction is considerably higher than in the uplink (i.e., $D_\beta = 6.89$ s, $D_\alpha = 41.18$ s) : this confirms that the packet generation rate in the downlink has a higher impact on the delay, with respect to its uplink counterpart. Another remarkable element is that downlink delay tends to be more homogeneous than uplink delay. Besides the visual representation of this behavior from the maps of Figure 6.11, we defined and calculated the delay variation coefficients σ_u^* and σ_d^*, for respectively the uplink and the downlink. These are defined as the ratio between the standard deviation and the average, and their computation for this scenario gives $\sigma_u^* = 0.495$ and $\sigma_d^* = 0.140$. The reason behind the higher homogeneity of the downlink direction can be found in the increased link congestion in the surroundings of the data collectors. Nodes in

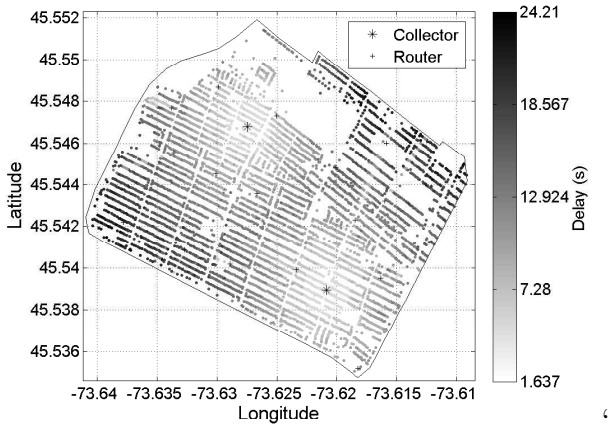
Table 6.2 95% confidence intervals of the average delay (expressed in seconds) for the different types of traffic.

	D_α (s)				D_β (s)			
$1/\lambda_d$	$1/\lambda_u$ (hours)				$1/\lambda_u$ (hours)			
(hours)	1	0.5	0.25	0.125	1	0.5	0.25	0.125
4.0	**6.48**±0.03	6.57±0.03	6.71±0.07	6.90±0.09	**6.79**±0.16	6.92±0.13	7.18±0.14	7.47±0.17
3.0	6.49±0.03	6.58±0.03	6.73±0.06	6.92±0.10	6.91±0.15	7.08±0.16	7.41±0.27	7.69±0.22
2.0	6.52±0.03	6.61±0.03	6.76±0.07	6.94±0.13	7.33±0.24	7.53±0.27	7.90±0.34	8.29±0.41
1.0	6.59±0.03	6.68±0.05	6.84±0.08	7.05±0.13	10.75±0.77	11.73±1.42	13.19±1.54	15.52±2.17
0.5	6.64±0.04	6.73±0.06	6.89±0.08	**7.10**±0.13	34.89±1.48	37.24±1.89	41.18±2.34	**45.16**±3.21
	D_α (s) - with broadcast traffic				D_β (s) - with broadcast traffic			
$1/\lambda_d$	$1/\lambda_u$ (hours)				$1/\lambda_u$ (hours)			
(hours)	1	0.5	0.25	0.125	1	0.5	0.25	0.125
4.0	**6.48**±0.02	6.57±0.04	6.71±0.06	6.90±0.08	**7.62**±0.17	7.81±0.16	8.15±0.20	8.51±0.24
3.0	6.50±0.02	6.58±0.04	6.74±0.06	6.93±0.10	7.78±0.19	8.00±0.22	8.35±0.16	8.74±0.22
2.0	6.52±0.03	6.61±0.05	6.76±0.06	6.94±0.10	8.19±0.28	8.43±0.24	8.85±0.29	9.45±0.46
1.0	6.59±0.03	6.68±0.04	6.84±0.07	7.05±0.13	11.64±0.85	12.67±1.31	14.27±1.23	16.50±1.87
0.5	6.64±0.04	6.73±0.06	6.91±0.07	**7.11**±0.15	35.49±1.57	37.97±1.90	42.54±2.33	**46.90**±3.68

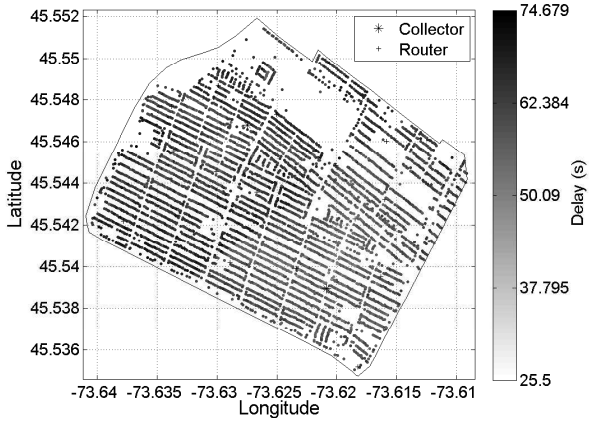


(a) D_α (b) D_β

Figure 6.10 Confidence intervals for D_α and D_β traffic without broadcast traffic.



(a) D_α



(b) D_β

Figure 6.11 Heat-map of the delay in the scenario with one broadcast transmission per day, $1/\lambda_d = 0.5$ h and $1/\lambda_u = 0.25$ h.

these areas suffer from the proximity to the collector and its limited buffer size, that causes additional waiting times to the packet transmission.

Continuing the analysis, we investigated the impact of broadcast transmissions (see lower part of Table 6.2) on the delay of the α and β traffic. We can notice that the uplink communications are not sensitively affected by the introduction of a broadcast transmission : the interval of D_α spans from 6.48 s to 7.11 s (whereas, the variation range without broadcast was from 6.48 s to 7.10 s. On the other hand, a small increase in D_β is observed : suffice it to compare its average values in the case with broadcast (i.e., from 7.62 s to 46.90 s) to the corresponding average values without the broadcast (i.e., from 6.79 s to 45.16 s). D_γ resulted to be considerably higher than D_α and D_β, with average values upper than 5 minutes.

6.4.4 Activity time percentage

Another sensitive index in the context of smart meter performance analysis is the activity time percentage $\chi(i)$, which represents the fraction of the simulated time in which node i is actually transmitting a packet. The importance of this parameter is driven by the recent concerns about the electromagnetic impact on public health, issued by the massive installation of smart meters in residential premises. The average activity time χ of all nodes can be computed as follows :

$$\chi = \frac{\sum_{\forall i \in V} \chi(i)}{|V|} = \frac{\sum_{\forall i \in V} N_t(i)}{T_s \, |V|} \tag{6.4}$$

Since χ takes into account all the nodes in the network, and since the main concerns are related to the smart meters, we decided to define 3 additional indexes, χ_m, χ_r, and χ_c, representing the average activity time percentage of respectively smart meters, routers, and data collectors, which are computed as follows :

$$\chi_m = \frac{\sum_{\forall i \in M} \chi(i)}{|M|} = \frac{\sum_{\forall i \in M} N_t(i)}{T_s \, |M|} \tag{6.5}$$

$$\chi_r = \frac{\sum_{\forall i \in R} \chi(i)}{|R|} = \frac{\sum_{\forall i \in R} N_t(i)}{T_s \, |R|} \tag{6.6}$$

$$\chi_c = \frac{\sum_{\forall i \in C} \chi(i)}{|C|} = \frac{\sum_{\forall i \in C} N_t(i)}{T_s \, |C|} \tag{6.7}$$

95% confidence intervals for χ_m, χ_r and χ_c are reported in Table 6.3. We can notice that, as expected, the activity time percentage increases as the packet generation rates increase : we observe that the variation of the activity time percentage with respect to the variation of $1/\lambda_d$ is much more evident than with respect to the variation of $1/\lambda_u$. However, numerical values of χ_m are always between 0.103 and 0.522 % : this is interesting because it can be used to reply to the public concerns about an excessive human exposure to wireless waves produced by smart meters. In fact, our results show that this kind of devices were in transmission mode for only a very low percentage of time, much lower than many others wireless devices used in everyday life, such as smartphones, and cordless phones. Particularly interesting is also the mean activity time of data collectors : in all the scenarios with $1/\lambda_d = 0.5$ h, the data collectors are active for 50% of the time, on average.

Table 6.3 Average activity time percentages for smart meters (χ_m), routers (χ_r) and data collectors (χ_c) for the different type of traffic according to different levels of $1/\lambda_u$ (rows) and $1/\lambda_d$ (columns) without broadcast style traffic.

	χ_m (%)			
$1/\lambda_d$	$1/\lambda_u$ (hours)			
(hours)	1	0.5	0.25	0.125
4.0	**1.03e-1±3.63e-4**	1.87e-1±1.04e-3	3.29e-1±3.66e-3	4.59e-1±7.94e-3
3.0	1.08e-1±3.89e-4	1.93e-1±1.04e-3	3.34e-1±3.48e-3	4.65e-1±8.88e-3
2.0	1.18e-1±4.61e-4	2.02e-1±1.15e-3	3.44e-1±4.11e-3	4.75e-1±1.13e-2
1.0	1.39e-1±4.98e-4	2.25e-1±1.58e-3	3.69e-1±4.36e-3	5.02e-1±1.13e-2
0.5	1.60e-1±7.95e-4	2.45e-1±1.68e-3	3.89e-1±4.20e-3	**5.22e-1±8.48e-3**
	χ_r (%)			
$1/\lambda_d$	$1/\lambda_u$ (hours)			
(hours)	1	0.5	0.25	0.125
4.0	1.08±7.19e-3	1.30±1.20e-2	1.68±4.03e-2	2.04±7.86e-2
3.0	1.34±8.63e-3	1.58±1.63e-2	1.97±3.94e-2	2.35±1.07e-1
2.0	1.84±1.75e-2	2.08±2.21e-2	2.50±4.81e-2	2.89±9.85e-2
1.0	3.03±3.26e-2	3.29±5.41e-2	3.73±6.53e-2	4.17±1.36e-1
0.5	3.57±7.45e-3	3.78±1.89e-2	4.14±3.46e-2	**4.50±7.46e-2**
	Activity time of data collectors : χ_c(%)			
$1/\lambda_d$	$1/\lambda_u$ (hours)			
(hours)	1	0.5	0.25	0.125
4.0	14.8±8.94e-2	15.1±1.60e-1	15.6±2.37e-1	16.1±3.38e-1
3.0	19.3±1.65e-1	19.7±1.97e-1	20.3±4.07e-1	21.0±4.80e-1
2.0	27.5±3.06e-1	28.1±2.88e-1	29.0±4.83e-1	30.0±7.53e-1
1.0	45.7±5.27e-1	46.2±9.75e-1	47.2±7.94e-1	47.9±7.68e-1
0.5	**50±5.97e-2**	**50±5.88e-2**	**50±1.75e-1**	**50±3.78e-2**

6.4.5 Computational time

RF-mesh systems are usually large-scale with thousands of nodes, therefore the computational burden can be a serious impairment to simulation and needs to be carefully taken into account. With the presented simulation framework, the average time to perform a simulation of 6051-node instances over a simulated time of 7 days was of around 30 minutes, on a machine with a *AMD A8-4500M CPU* working at 1.90 GHz. This is a very important result because it proves the capacity to simulate large-scale systems in a considerably short time.

6.4.6 Scalability analysis

Another important point in the evaluation of a large-scale network simulator is the variation of the computational time, as the size of the network increases. To this purpose, 12 topologies were created with a number of nodes ranging from 1718 to 20798. The topologies were used to perform $7 - day$ simulations, whose computational times are reported in Figure 6.12. A quasi-linear growth of the computational tie is observed. We computed the quadratic regression of the computational times with respect to the number of nodes and the result was the following 2^{nd} degree polynomial :

$$y = 7.469 \cdot 10^{-5}x^2 + 2.746 \cdot 10^{-3}x - 103.802 \tag{6.8}$$

The very low 2^{nd} degree exponent explains the low concavity of the quadratic regression curve and the quasi-linear behavior of the computational time with respect to the number of nodes„ in the analyzed domain (between 1718 and 20798).

6.5 Conclusion

As smart grids proliferate around the world and AMI systems are widely deployed, more and more applications will likely be introduced. Therefore, it is essential for the power utilities to have appropriate tools to be able to assess the performance of the system and the impact of new applications on it. In this paper, we have presented a framework to assess the performance of AMI systems that removes some of the limitations of previous work related to AMI performance, such as scalability or adaptability.

The framework yielded an effective simulation tool that allows for great flexibility in the selection of system features (e.g., the area under consideration, the number of nodes). The implemented topology generation tool grants the possibility to create customized instances, and allows both to manually choose the location of routers and collectors, and to be combined to an optimization tool for the location of routers and data collectors. The observed computational efficiency permits the analysis of instances of thousands of nodes in a very reasonable time.

The numerical results used to illustrate the framework capabilities have nonetheless brought to light some useful insights into the AMI system in a densely populated NAN. There is a performance degradation in the surroundings of the collectors, where multiple traffic streams coexist and mutually disturb each other. Difficulties in the handling of broadcast traffic were

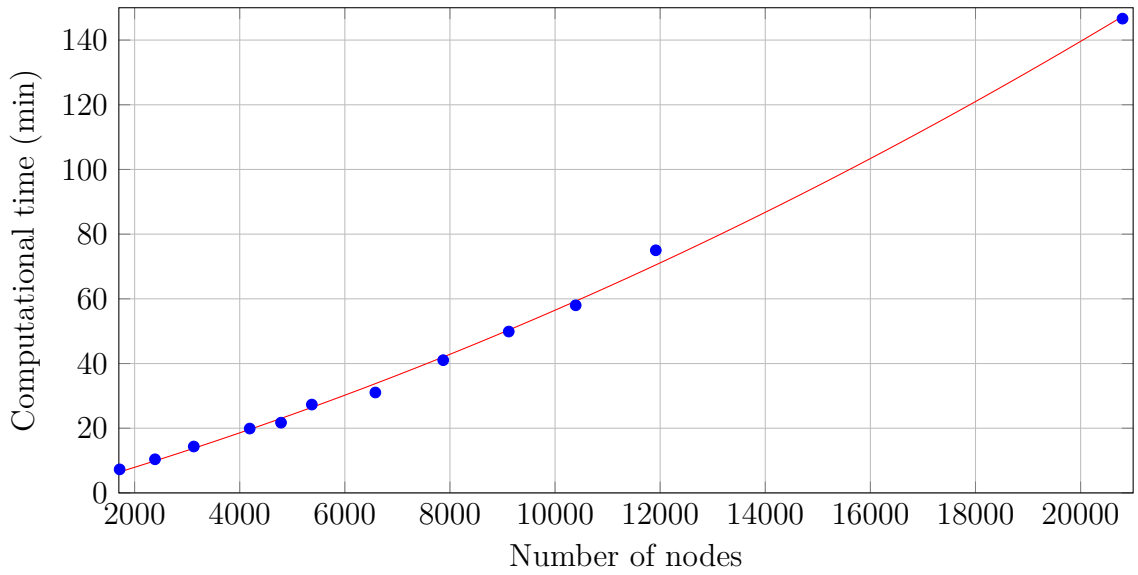


Figure 6.12 Computational time to perform a 7-day simulation using topologies with a number of nodes ranging between 1718 and 20798. The quadratic regression is reported in red.

observed, but the uplink delay is not particularly affected by the broadcast traffic. The 95 % confidence interval analysis was included to show the accuracy of the simulation results. It would be interesting to compare those findings with other types of AMI deployment.

CHAPTER 7 ARTICLE 4 : A MARKOV-MODULATED END-TO-END DELAY ANALYSIS OF LARGE-SCALE RF-MESH NETWORKS WITH TIME-SLOTTED ALOHA AND FHSS

F. Malandra et B. Sansò, (2016), "A Markov-modulated End-to-end Delay Analysis of Large-scale RF-Mesh Networks with Time-slotted ALOHA and FHSS", IEEE, Transactions on Wireless Communications, submitted to.

Abstract

A new mathematical model and a methodology are proposed to evaluate the performance of large-scale RF-Mesh Networks that use time-slotted ALOHA with Frequency Hopping Spread Spectrum. An analytic formulation for the delay, based on Markov-modulated modelling of the system, is derived. The formula can be extended to evaluate other important performance metrics. The proposed methodology is applied to a large scale network of several thousands of nodes, and numerical results are reported to show the wide variety of performance evaluations that are enabled. The usefulness of the assessment of the feasibility of different types of applications (e.g., smart-metering, sensor networks, IoT) is shown. An analysis of the scalability of this methodology and a comparison with simulation results are also presented.

7.1 Introduction

Wireless Mesh Networks (WMNs) are dynamically self-configured and self-organized networks. They are increasingly widespread because of their low complexity, their ease of installation, and their robustness. WMNs are nowadays used as a transmission support for a wide variety of new applications, such as the Internet of Things (IoT) and Smart Grid, which are characterized by a large number of simple and low cost devices that need to communicate with a central element of the architecture.

AMIs are often used to enable Smart Grid applications. Their main objective is to connect the large number of residential and commercial smart meters to the MDMS of the power utility; however, due to the high number of emerging applications, AMIs and their communication infrastructures can be used for multiple applications, different from that they were designed for. For instance, AMIs were originally conceived for the remote reading of the electricity consumption by the power utility, but are currently being considered as a trans-

mission support for load management and demand-response applications. The integration of different applications, each characterized by its own communication traffic, may not be feasible in a certain network; the feasibility of the integration needs to be assessed by means of thorough performance studies of the network.

The communication of a large number of devices in a densely populated wireless environment, such as an AMI, can easily provoke interference, with consequences on the overall network performance. The effect of interference on performance is usually difficult to evaluate because of the complex nature of the system, usually characterized by a high number of nodes. In fact, the most common approach in interference analysis is to use an interference matrix, whose complexity grows quadratically with the number of nodes. Another peculiarity of many Smart Grid applications is that they require bidirectional communications between the mesh nodes and the gateway and have constraints on the latency in packet reception. The estimation of the end-to-end delay for each origin-destination pair in a large-scale WMN is difficult because many factors affect the delay a packet can experience (e.g., MAC algorithm, interference, routing link congestion).

Many different WMN technologies are currently used in AMIs, such as WiFi and LPWAN, but the so-called RF-Mesh system seems to be particularly popular. Suffice it to say that three of the most important smart meters manufacturers (Aclara, Itron, and Landys&Gyr, that own 76.5% of the market share in North America, according to Navigant Research (2016)) equip their smart meters with RF-Mesh radios[1]. RF-Mesh technology uses the free and unlicensed ISM bandwidth to create a mesh topology connecting the smart meters to a data collector in an AMI. RF-Mesh technology prevails over the other solutions thanks to its reduced equipment cost, its proprietary infrastructure, and its ease of installation. However, the performance analysis for RF-Mesh AMIs is even more difficult than for other solutions because it is less standardized, its implementation details are often covered by harsh confidentiality agreements (e.g., on routing, number of frequency channels), and very little data are currently publicly available. As a result, the literature on RF-Mesh performance is not as complete as that focusing on other kinds of WMNs (e.g., WiFi).

This paper aims at contributing to the knowledge in this field by proposing a framework for the performance analysis of a RF-Mesh-based AMI, based on the packet collision probability and on the delay. The RF-Mesh system is modeled as a MMS, which allows to account for multiple complex implementation details, that - to our knowledge - have not been considered simultaneously in the existing body of literature (e.g., buffer size, retransmission probability).

[1]For more information, `http://www.greentechmedia.com/articles/read/who-are-the-top-ten-vendors-in-smart-grid/all`

The rest of the document is structured as follows : Section 7.2 reports a literature review on performance studies of WMNs, with particular attention to RF-Mesh systems; Section 7.3 contains the description of the communication system under study; Section 7.4 deals with the MMS model of the system; Section 7.5 presents the performance analysis framework and some numerical results; the conclusions of the study are discussed in Section 7.6.

7.2 State of the art

An increasing deal of research proposes performance studies for wireless AMIs. More details on the recent initiatives and the ongoing projects involving AMIs can be found in Abdulla (2015); Foudeh and Mokhtar (2015); Garcia-Hernandez (2015); Renofio et al. (2016), whereas recent interesting surveys on AMI performance analysis can be found in Mohassel et al. (2014); Namboodiri et al. (2012); Sharma and Saini (2015). The existing body of literature proposes performance studies focusing on several specific aspects of AMIs, such as security, resiliency, physical and network layer protocols.

The theme of security is central in many studies (Anzalchi and Sarwat (2015); Bhatia and Bodade (2014); Somkaew et al. (2014); Soykan et al. (2015); Tonyali et al. (2016)) because AMIs usually deal with confidential data, that need to be protected. AMIs are prone to several types of attacks and physical layer security techniques, such as data encryption (e.g., in Somkaew et al. (2014); Soykan et al. (2015); Tonyali et al. (2016)) and strong authentication mechanisms (e.g., in Lee et al. (2014); Thomas et al. (2012)), are proposed as countermeasures.

Some other authors study the resiliency of AMI networks, evaluating their capacity to resist to unexpected node or link failures (e.g., in Renofio et al. (2016)). A large branch of the existing literature deals with the performance of the physical layer : the authors investigate new coding schemes (e.g., in Arias and Rodriguez (2015); Ayala and Rodriguez (2014); Le and Benjapolakul (2016)), new spectrum opportunities (e.g., in Khan et al. (2016); Parvez et al. (2016); Riascos et al. (2015); Sousa et al. (2015)), and new modulation techniques (e.g., in Oh et al. (2015); Zeng et al. (2012)) to improve AMI performance. Other authors analyze the performance at network level : a large multitude of routing protocols (e.g., RPL, AODV, layer-based) are thoroughly discussed (Elyengui et al. (2015); Hu et al. (2015); Ramirez et al. (2015)). The large majority of the above-mentioned papers evaluate the performance of AMIs adopting wireless technologies that are considered standards, such as WiFi (e.g., in Hu et al. (2015); Michaloliakos et al. (2016); Ramirez et al. (2015); Sousa et al. (2015); Tonyali et al. (2016)), cellular networks (e.g., GSM in Das and Saikia (2015), 3G in Athanasios and Cottis (2016), and LTE in Carlesso et al. (2015); Yaacoub and Kadri (2015)), ZigBee (e.g., in Chang

et al. (2015); Chi et al. (2016); Parvez et al. (2016); Peng and Huang (2016); Somkaew et al. (2014)), Bluetooth (e.g., in Wang et al. (2013)), and LPWANs (e.g., in Vangelista et al. (2015)).

The three approaches have different peculiarities resulting, in some cases, complementary. For instance, a properly designed RF-Mesh simulator provides a good system performance overview, proving to be helpful for conceiving and testing new system features (e.g., the design of a new routing protocol). The flexibility in the simulation is unfortunately balanced by a computational burden that might limit a large scalability. Field trials provide useful insights into the performance of already implemented RF-Mesh systems. The use of field-trials permits to identify abnormal behaviors (e.g., unexpected bottlenecks) which are difficult to discover in a simulated environment. However, field trials are expensive and difficult to be put in place : a large number of devices needs to be installed and this might be unfeasible or lead to unaffordable costs for the research community. Finally, the analytic approach reduces the computational cost and allows to generalize the findings to different instances and use cases; however, it requires preliminary assumptions that might reduce the fitness to actually implemented systems.

As will be detailed in Section 7.4, an analytic approach is followed in this paper in order to calculate the average delay in the network and, based on that, to carry out a detailed performance analysis of RF-Mesh AMIs. In order to have a model that better suits the dynamics of the RF-Mesh system at hand, a MMS model was chosen. MMSs are extensively used in the wireless domain to model the traffic generation (e.g., in Andronov and Vishnevsky (2016)) and the varying propagating conditions (e.g., in Zheng et al. (2013)). As a consequence, many authors exploited this complex and powerful mathematical tool to carry out performance evaluations on wireless networks. In this paper, a MMS is used to model a multi-hop RF-Mesh AMI with time-slotted ALOHA (hereafter referred to as s-ALOHA) and FHSS.

Although some work on the performance of systems with s-ALOHA exists (e.g., Masood et al. (2015); Narvare and Shrivastava (2014); Rom and Sidi (2012)), in no case, the FHSS or the multi-hop nature of mesh networks are considered. To the best of our knowledge this is the first time that a comprehensive analytical study is carried out for s-ALOHA, multi-hop networks with FHSS.

7.3 Description of the system

7.3.1 Architecture

The system at hand is a WMN, composed of a set N of mesh nodes, a set R of mesh routers, and a set G of gateways. The total number of nodes is expressed by the set $V = N \cup R \cup G$. In the rest of the document, we will focus on a RF-Mesh AMI, in which N represents the smart meters, G represents the data collectors, and R the routers (see Figure 7.1). In order to define the links in the topology, two covering rays are assumed : one for routers and collectors, and one for smart meters. The values for the covering ray vary according to the considered scenario : for instance, the smart meter ray is in the order of few hundred meters in a urban scenario, whereas it can extend beyond 1 km in a rural area. Routers and collectors have typically larger covering rays, because they are equipped with better radios and are usually installed on higher sites (e.g., light poles). A link exists between two nodes if they are located at a relative distance lower than their covering rays. In the rest of the paper, we denote the set of neighbors of the i-th node with $\mathcal{N}(i)$.

7.3.2 The physical layer

Packet transmissions take place in the ISM bandwidth centered at 900 MHz, which is shared with a large multitude of other devices. It is well known that wireless transmissions may collide, leading to incorrect packet receptions and consequent retransmissions. This, in turn, increases network delay and reduces system performance. If a single wireless channel is available, a collision occurs in the transmission from node i to node j when at least one node $k \in \mathcal{N}(j)$ transmits simultaneously. The FHSS protocol is usually adopted in order to mitigate interference : it is a spread spectrum technique that dynamically changes the transmission frequency, according to a predefined pseudo-random sequence, known to all the nodes in the network. With FHSS, a collision can occur only if node $k \in \mathcal{N}(j)$ transmits simultaneously in the same frequency channel used in transmission $i \rightarrow j$. In our model we denote by $\mathcal{F}$ the set of available frequency channels (80 channels of 300 kHz each in our model).

7.3.3 The MAC layer

The system at hand is characterized by a s-ALOHA access. The time is subdivided in time-slots of duration $\tau = 0.7$ s. At the beginning of each time-slot, a node that has packets in its buffer attempts a transmission with probability p_τ. Then, if a collision is observed in the same time-slot, the packet is assumed not to be correctly received and will be retransmitted

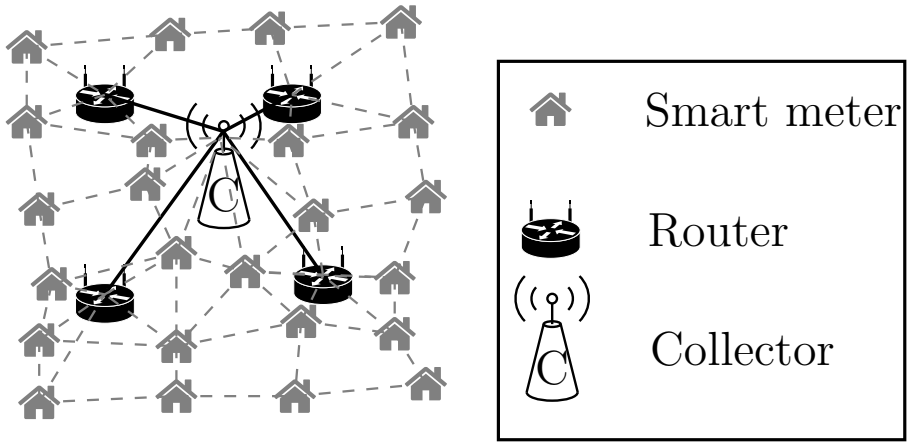


Figure 7.1 A scheme of the architecture of the RF-Mesh AMI under study.

in one of the subsequent time-slots with probability p_r. Although there are other possible causes that might prevent the correct reception of a packet, such as fading or the presence of an obstacle in the propagation, they are out of the scope of this research work.

7.3.4 Routing

Two different communication streams are considered in this model : one in downlink, from the data collectors to all the smart meters, and one in uplink, from all the smart meters to the data collectors. A shortest-path-based routing mechanism is adopted for both downlink and uplink traffic. The shortest paths from each mesh node to a single data collector are bidirectional. If we denote by $\Omega_u(i)$ the set of links in the uplink shortest path, $\Omega_d(i)$ (i.e., the downlink shortest path) contains the same but reversed links. Despite the use of a static routing might seem a too simplistic assumption in a dynamic wireless mesh context, it is a widely employed solution in smart grid communications studies (e.g., Akkaya et al. (2015); Ghasemkhani et al. (2016); Shao et al. (2015)), and it permits to derive closed-form expressions for the delay. Other relevant parameters are :

$\xi(i)$ the number of shortest paths containing node i as an intermediate node (neither source nor destination);

$\Gamma(i)$ a subset of $\mathcal{N}(i)$ including only nodes to which node i actually transmits, according to the shortest path routing[2];

$\alpha_i(j)$ the percentage of traffic at node i that will be transmitted to node $j : j \in \Gamma(i)$.

[2] $\Gamma(i) = \{j \in \mathcal{N}(i) : (i,j) \in \Omega_u(i) \vee (i,j) \in \Omega_d(j)\}$

7.3.5 Traffic characterization

Downlink and uplink traffic are assumed to be Poisson-distributed with mean values λ_d and λ_u, respectively. This assumption permits an easy aggregation of the traffic coming from different applications. As for a known property of Poisson-distributed processes, the aggregation of a multitude W of Poisson-distributed traffics, each with mean value λ_w, is also Poisson-distributed with mean value $\lambda = \sum_{w \in W} \lambda_w$. The mean transmission rate of node i, indicated with $\lambda(i)$, represents both the traffic generated by node i and the traffic flowing through the ξ_i shortest paths. Analytic expressions for $\lambda(i)$ can be found for the three different device types :

$$\lambda(i) = \begin{cases} \lambda_u + \xi_i(\lambda_u + \lambda_d) & \text{if } i \text{ is a smart meter} \\ \xi_i(\lambda_u + \lambda_d) & \text{if } i \text{ is a router} \\ |N|\lambda_d & \text{if } i \text{ is a data collector} \end{cases} \tag{7.1}$$

The number X_i of packets being generated or arriving at node i during a time-slot is:

$$P(X_i = k) = \frac{(\lambda(i)\tau)^k e^{-\lambda(i)\tau}}{k!} \tag{7.2}$$

Since low data rates are assumed, we neglect the possibility of having $X_i > 1$ and compute γ_i, the probability that node i produces or receives one packet, as follows :

$$\gamma_i = P(X_i = 1) = \lambda(i)\tau e^{-\lambda(i)\tau} \tag{7.3}$$

7.3.6 Example of system

A toy topology composed of 3 smart meters (s_1, s_2 and s_3), 2 routers (r_1 and r_2) and the data collector c is reported in Figure 7.2 in order to better illustrate the routing and the traffic characterization that is carried out in this work. In the graph, solid bidirectional arrows indicate the presence of links between pairs of nodes, dashed directed arrows represent downlink traffic, whereas directed dashed-dotted arrows represent uplink traffic. The labels on directed arrows represent the mean values of the aggregated traffic flowing on that link.

In the uplink, each smart meter produces packets at a rate of λ_u. s_1 and s_2 transmit at a rate λ_u to respectively r_1 and r_2. The two routers forward the same traffic to node s_3. Node s_3 receives $2\lambda_u$ from the routers, and generates packets at a rate λ_u; therefore, s_3 transmits $3\lambda_u$ to the data collector which is the sink for the uplink traffic. In the downlink, the data

collector c generates packets at a rate $3\lambda_d$. Accordingly, we have $3\lambda_d$ on the link (c, s_3), λ_d on the links (s_3, r_1), (s_3, r_2), (r_1, s_1), and (r_2, s_2).

In what follows, $\Gamma(i)$ and $\alpha_i(j)$ are computed :

$$\begin{array}{ll}
\Gamma(s_1) = \{r_1\} & \alpha_{s_1}(r_1) = 1 \\
\Gamma(s_2) = \{r_2\} & \alpha_{s_2}(r_2) = 1 \\
\Gamma(s_3) = \{r_1, r_2, c\} & \alpha_{s_3}(r_1) = \frac{\lambda_d}{2\lambda_d+3\lambda_u},\ \alpha_{s_3}(r_2) = \frac{\lambda_d}{2\lambda_d+3\lambda_u},\ \alpha_{s_3}(c) = \frac{3\lambda_u}{2\lambda_d+3\lambda_u} \\
\Gamma(r_1) = \{s_1, s_3\} & \alpha_{r_1}(s_1) = \frac{\lambda_d}{\lambda_d+\lambda_u},\ \alpha_{r_1}(s_3) = \frac{\lambda_u}{\lambda_d+\lambda_u} \\
\Gamma(r_2) = \{s_2, s_3\} & \alpha_{r_2}(s_2) = \frac{\lambda_d}{\lambda_d+\lambda_u},\ \alpha_{r_2}(s_3) = \frac{\lambda_u}{\lambda_d+\lambda_u} \\
\Gamma(c) = \{s_3\} & \alpha_c(s_3) = 1
\end{array}$$

We can infer that each node i has a different transmission rate $\lambda(i)$ that depends on the nominal transmission rates in the uplink and downlink, and also on ξ_i, defined as the number of shortest paths that contain i as an intermediate node.

7.4 Markov-modulated System

A MMS is a Markov system in which the transition probabilities are not fixed, but depend on the evolution of the state of the system. In our case, the state of node i at time-slot t is subdivided into three *sequential* phases : $S_B(i,t)$ (Beginning), $S_T(i,t)$ (Transmission), $S_E(i,t)$ (End). In particular, each node will go through the 3 phases at each time-slot. Mathematically, if we denote by π_S the probability of being in a state S, we have that :

$$\sum_{S \in S_B(i,t)} \pi_S = 1, \sum_{S \in S_T(i,t)} \pi_S = 1, \sum_{S \in S_E(i,t)} \pi_S = 1 \qquad \forall i \in V, \forall t \tag{7.4}$$

In Figure 7.3, a scheme with $S_B(i,t)$, $S_T(i,t)$, $S_E(i,t)$, and $S_B(i,t+1)$ is reported to illustrate the sequence of 4 consecutive states.

7.4.1 Beginning of the time-slot

In the *beginning phase* $S_B(i,t)$, the state of node i is represented by the number of packets in its buffer at the beginning of time-slot t. The buffer size, denoted with Z, is assumed to be the same for all the nodes in the system and constant over time. Accordingly, node i

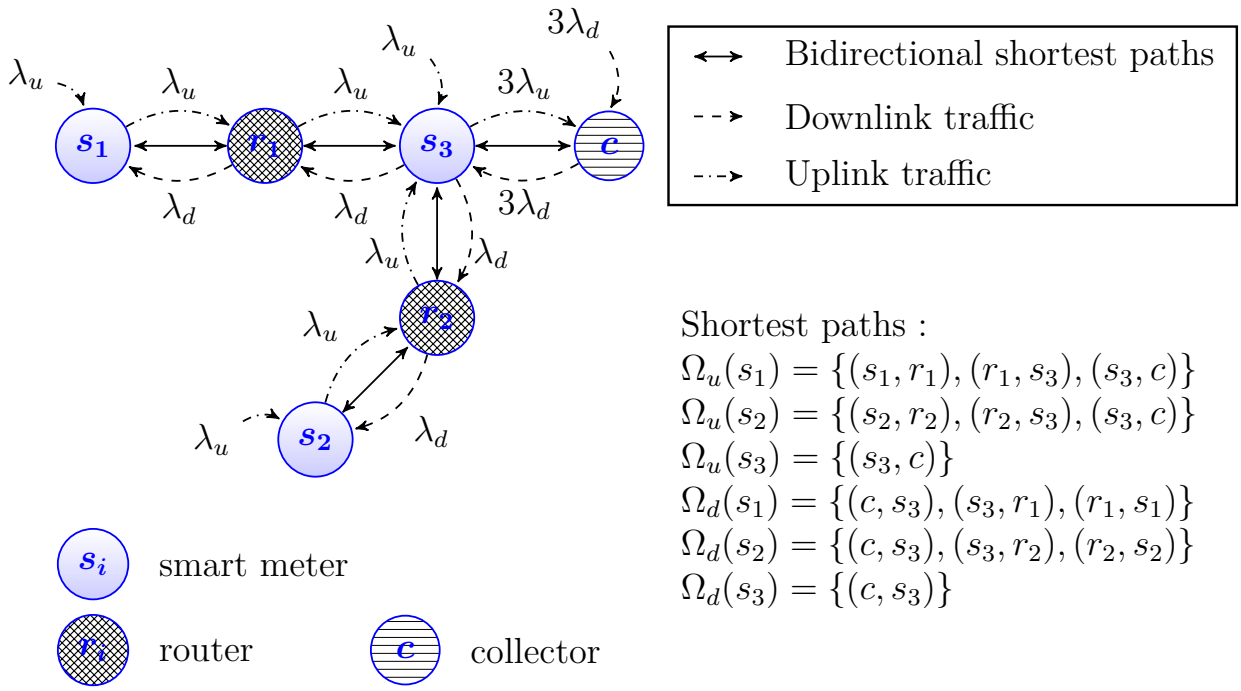


Figure 7.2 A toy topology with 3 smart meters (s_1, s_2 and s_3), 2 routers (r_1 and r_2) and the data collector c.

can be in $Z + 1$ states at the *beginning* of a time-slot. The probability that node i is in a certain state — hereafter called state probability — with n packets in the buffer is $\pi_B(i, n, t)$. Condition (7.4) for the beginning phase can be rewritten as $\sum_{n=0}^{Z} \pi_B(i, n, t) = 1 \quad \forall i \in V, \forall t$. For the subsequent analysis, it is convenient to use a vectorial notation. Therefore, let us define the *beginning* state probability vector $\mathbf{\Pi}_B(t)$ at time-slot t as follows :

$$\mathbf{\Pi}_B(t) = \{\pi_B(i, n, t) : \quad \forall i \in V, \forall n \in [0, Z]\} \tag{7.5}$$

7.4.2 Transmission phase

During the transmission phase, node i can be either idle or transmitting to any node in the set $\Gamma(i)$ (see Section 7.3.4). Therefore, the states in the *transmission phase* $S_T(i, t)$, illustrated in Figure 7.3, are characterized by two indexes : the first is either $\varnothing$, when node i is idle, or $\Gamma_k(i)$, when it is transmitting to the k-th node of set $\Gamma(i)$; the second represents the number of packets in the buffer. The state probabilities in the *transmission* phase are defined as follows :

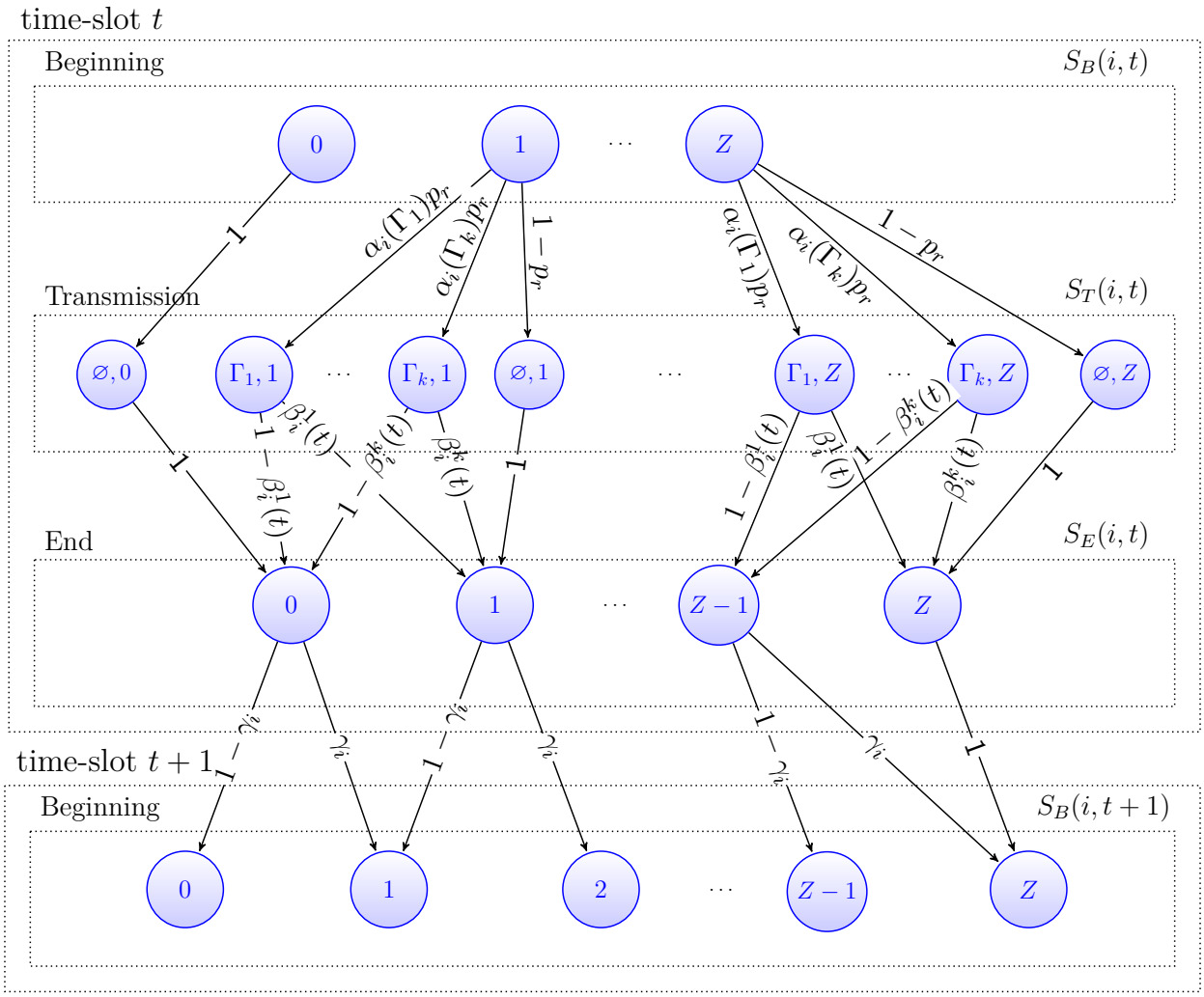


Figure 7.3 The Markov chain representing the evolution of the states of node i from time-slot t (i.e., $S_B(i,t)$, $S_T(i,t)$, and $S_E(i,t)$) to time-slot $t+1$(i.e., $S_B(i,t+1)$).

- $\pi_T(i, \varnothing, n, t)$, the probability of being idle at time-slot t with n packets in the buffer during the transmission phase of time-slot t
- $\pi_T(i, \Gamma_k(i), n, t)$, the probability of transmitting a packet from node i to node $\Gamma_k(i)$ with n packets in the buffer of node i during the transmission phase of time-slot t

Note that in Figure 7.3, the compact notation Γ_k was used instead of $\Gamma_k(i)$, and the indication of the node and of the time-slot are omitted : therefore, the transmitting state $S_T(i, \Gamma_k(i), n, t)$ is replaced by the more compact form (Γ_k, n).

The transition probabilities between the *beginning* phase and the *transmission* phase depend on the two parameters $\alpha_i(\Gamma_k)$ (i.e., the portion of traffic of node i transmitted to node Γ_k, see Section 7.3.4) and p_r (i.e., the probability to transmit or retransmit, see Section 7.3.3), and can be computed as follows :

$$\pi_T(i, \varnothing, 0, t) = \pi_B(i, 0, t) \tag{7.6}$$

$$\pi_T(i, \Gamma_k, n, t) = \alpha_i(\Gamma_k) p_r \pi_B(i, n, t) \tag{7.7}$$

$$\pi_T(i, \varnothing, n, t) = (1 - p_r) \pi_B(i, n, t) \tag{7.8}$$

Eq. (7.6) means that if node i has an empty buffer, it will stay idle with probability equal to one; Eq. (7.7) says that node i passes from state $S_B(i, n, t)$ to state $S_T(i, \Gamma_k, n, t)$ with probability $\alpha_i(\Gamma_k) p_r$ (i.e., the probability to transmit to node Γ_k); whereas Eq. (7.8) states that node i passes from state $S_B(i, n, t)$ to $S_T(i, \varnothing, n, t)$ with probability $1 - p_r$ (i.e., the probability not to transmit). It can be easily proven that condition (7.4) is verified :

$$\sum_{n=0}^{Z} \Pi(i, \varnothing, n, t) + \sum_{n=1}^{Z} \sum_{\Gamma_k \in \Gamma(i)} \Pi(i, \Gamma_k, n, t) = 1 \quad \forall i \in V, \forall t \tag{7.9}$$

The probability that node i stays idle or transmits to one of its neighbours Γ_k can be grouped as :

$$\pi_T(i, \varnothing, t) = \sum_{n=0}^{Z} \pi_T(i, \varnothing, n, t) \tag{7.10}$$

$$\pi_T(i, \Gamma_k, t) = \sum_{n=0}^{Z} \pi_T(i, \Gamma_k, n, t) \tag{7.11}$$

Also, the probability that node i transmits during time-slot t is :

$$\pi_T(i,t) = \sum_{\Gamma_k \in \Gamma(i)} \pi_T(i,\Gamma,t) = 1 - \pi_T(i,\varnothing,t) \tag{7.12}$$

The *transmission* state probability vector $\mathbf{\Pi}_T(t)$ at time-slot t is defined as follows :

$$\mathbf{\Pi}_T(t) = \{\pi_T(i,\varnothing,n,t), \pi_T(i,\Gamma_k,n,t) : \quad \forall i \in V, \forall \Gamma_k \in \Gamma(i), \forall n \in [0,Z]\} \tag{7.13}$$

7.4.3 End of the time-slot

The state at the end of the time-slot is represented by the number of packets in the buffer. As for the *beginning* phase, we have $Z+1$ possible states in $S_E(i,t)$ for each node i at each time-slot t. The probability that node i has n packets in the buffer at the end of time-slot t is indicated with $\pi_E(i,n,t)$, given that $\sum_{n=0}^{Z} \pi_E(i,n,t) = 1 \quad \forall i \in V, \forall t$.

The transition from the *transmission* states to the end states depends on the collision probability in the transmission from node i to node $\Gamma_k(i)$, indicated with $\beta_i^{\Gamma_k}(t)$. Hereafter, for simplicity reasons, we will indicate $\beta_i^{\Gamma_k(i)}(t)$ by $\beta_i^k(t)$. We proposed an expression for the collision probability in a s-ALOHA system with FHSS and $|\mathcal{F}|$ available channels in Malandra and Sansò (2015). For the current model, we modified that expression as follows :

$$\beta_i^k(t) = 1 - \prod_{j \in \mathcal{N}(\Gamma_k) - \{i\}} \left(1 - \frac{1}{|\mathcal{F}|}\pi_T(j,t)\right) \tag{7.14}$$

The main difference is in term $\pi_T(j,t)$, defined in Section 7.4.2, which takes into account implementation details not considered in Malandra and Sansò (2015), such as retransmission probability and buffer size. Eq. (7.14) shows that a correct reception is observed in the communication $i \to \Gamma_k$ only if no interferer $j \in \mathcal{N}(\Gamma_k)$ is transmitting in the same frequency channel.

The transition probabilities are then computed as follows :

$$\pi_E(i,0,t) = \pi_T(i,\varnothing,0,t) + \sum_{\Gamma_k \in \Gamma(i)} \pi_T(i,\Gamma_k,1,t)(1-\beta_i^k(t)) \tag{7.15}$$

$$\pi_E(i,n,t) = \pi_T(i,\varnothing,n,t) + \sum_{\Gamma_k \in \Gamma(i)} \pi_T(i,\Gamma_k,n+1,t)(1-\beta_i^k(t)) + \\ + \sum_{\Gamma_k \in \Gamma(i)} \pi_T(i,\Gamma_k,l,t)\beta_i^k(t) \quad \forall n = 1,\ldots,Z-1 \tag{7.16}$$

$$\pi_E(i,Z,t) = \pi_T(i,\varnothing,Z,t) + \sum_{\Gamma_k \in \Gamma(i)} \pi_T(i,\Gamma_k,Z,t)\beta_i^k(t) \tag{7.17}$$

Eq. (7.15) states that node i has an empty buffer either if it had an empty buffer in the transmission phase or if it had one packet and transmitted without a collision; Eq. (7.16) says that node i has n packets in the buffer if it had n packets and stayed idle, if it had $n+1$ packets and transmitted without collision, and if it had n packets and transmitted with a collision; whereas Eq. (7.17) says that node i has Z packets at the end of the time-slot if it had Z packets in the transmission phase and either it was idle or it transmitted with a collision.

The *end* probability vector $\mathbf{\Pi}_E$ and the collision probability vector $\boldsymbol{\beta}(t)$ can be defined as follows :

$$\mathbf{\Pi}_E(t) = \{\pi_E(i,n,t) : \quad \forall i \in V, \forall n \in [0,Z]\} \tag{7.18}$$

$$\boldsymbol{\beta}(t) = \{\beta_i^k(t), \forall i \in V, \forall \Gamma_k \in \Gamma(i)\} \tag{7.19}$$

7.4.4 Time evolution

Between the end of time-slot t and the beginning of time-slot $t+1$, node i receives or produces a packet with probability γ_i, computed according to Eq. (7.3); it will maintain the same number of packets in the buffer with probability $1-\gamma_i$. An exception occurs when the buffer is full (Z packets) : in this situation, a new packet, generated or received, is lost with probability $\pi_E(i,Z,t)\gamma_i$. Mathematically :

$$\pi_B(i,0,t+1) = \pi_E(i,0,t)(1-\gamma_i) \tag{7.20}$$

$$\pi_B(i,n,t+1) = \pi_E(i,n,t)(1-\gamma_i) + \pi_E(i,n-1,t)\gamma_i, \ \ \forall n \in [1,z-1] \tag{7.21}$$

$$\pi_B(i,Z,t+1) = \pi_E(i,Z,t) + \gamma_i \pi_E(i,Z-1,t) \tag{7.22}$$

7.4.5 Stationary probabilities

Using a vector notation, we can express the transitions of the Markov state as $\mathbf{\Pi}_T(t) = f\Big(\mathbf{\Pi}_B(t)\Big)$, $\mathbf{\Pi}_E(t) = g\Big(\mathbf{\Pi}_T(t)\Big)$, and $\mathbf{\Pi}_B(t+1) = h\Big(\mathbf{\Pi}_E(t)\Big)$. The equations included in function f correspond to equations (7.6)-(7.8); those in function g to (7.15)-(7.17); those in function h to (7.20)-(7.22). One of the objectives of the implemented Markov-modulated model is to find stationary solutions, for which :

$$\mathbf{\Pi}_B(t+1) = h\Big(\mathbf{\Pi}_E(t)\Big) = h\Big(g\Big(\mathbf{\Pi}_T(t)\Big)\Big) = h\Big(g\Big(f\Big(\mathbf{\Pi}_B(t)\Big)\Big)\Big) = (h \circ g \circ f)\,(\mathbf{\Pi}_B(t)) \quad (7.23)$$

The high number of equations included in f, g, and h, as well as the unavoidable numerical errors, make it hard to obtain a perfect equality. As a consequence, a norm of the variation is defined as :

$$\Delta(t) = ||\mathbf{\Pi}_B(t) - \mathbf{\Pi}_B(t-1)|| = \sum_{i \in V} \sum_{n=0}^{Z} ||\pi_B(i,n,t) - \pi_B(i,n,t-1)|| \quad (7.24)$$

The *beginning* state probability vector $\mathbf{\Pi}_B(t^*)$ at time t^* is considered *stationary* if and only if $\Delta(t^*) < \epsilon$, ϵ being the chosen tolerance. At the equilibrium, we have :

$$\mathbf{\Pi}_B^* = \mathbf{\Pi}_B(t^*), \mathbf{\Pi}_T^* = \mathbf{\Pi}_T(t^*),\ \mathbf{\Pi}_E^* = \mathbf{\Pi}_E(t^*), \boldsymbol{\beta} = \boldsymbol{\beta}^*(t^*) \quad (7.25)$$

Note that vector $\mathbf{\Pi}_B$ has a number of elements equal to $(Z+1)|V|$ (e.g., it is 120000 in a topology with 6000 nodes and a buffer size $Z = 20$).

7.4.6 Network Delay

The stationary probabilities computed as per section 7.4.5 are used to find the average delays, which are fundamental, for example, to assess the feasibility of smart grid applications in a certain network. As introduced in Section 7.3, each smart meter $i \in N$ is the source of a stream of traffic directed to the data collector and forwarded along the shortest path $\Omega_u(i)$; node i is also the destination of a downlink flow originated by the collector and routed along the shortest path $\Omega_d(i)$. Accordingly, we can define two types of delay for each $i \in N$, one per communication flow : let $\mathcal{D}_u(i)$ and $\mathcal{D}_d(i)$ be, respectively, the average multi-hop delay for uplink and downlink directions at node i. A standard practice for multi-hop systems is to calculate the delay of a path traversing multiple links as the sum of the average delays in

each link $j-k$, called δ_{jk}, included in $\Omega_u(i)$ and $\Omega_d(i)$:

$$\begin{aligned} \mathcal{D}_u(i) &= \sum_{(j,k)\in\Omega_u(i)} \delta_{jk} \\ \mathcal{D}_d(i) &= \sum_{(j,k)\in\Omega_d(i)} \delta_{jk} \end{aligned} \tag{7.26}$$

In order to find $\mathcal{D}_u(i)$ and $\mathcal{D}_d(i)$, we first need to calculate δ_{jk}, which has two main components : an average queuing delay q_j and an average transmission delay μ_{jk}. The transmission delay μ_{jk} is calculated as the average number of time-slots that are necessary for a successful transmission from j to k multiplied by the time slot duration τ. The number of attempts to get a success can be computed as a geometric distribution with success probability $p_r(1-\beta_j^k)$, since at each time-slot a successful transmission is achieved if the node transmits, which occurs with probability p_r, and if no collision is observed, which occurs with probability β_j^k. As a consequence, the transmission delay can be computed as :

$$\mu_{jk} = \frac{1}{p_r(1-\beta_j^k)}\tau \tag{7.27}$$

The average queueing delay at node j is the time necessary to transmit the average number of packets present in its buffer. It is computed as :

$$q_j = \sum_{k\in\Gamma(j)} \theta_j(k)\mu_{jk} \tag{7.28}$$

with θ_j as the average number of packets in the buffer of node j :

$$\theta_j = \sum_{n=0}^{Z} n\pi_B(j,n) \tag{7.29}$$

and with $\theta_j(k)$ as the fraction of θ_j directed to node k :

$$\theta_j(k) = \alpha_j(k)\theta_j \quad \forall k \in \Gamma(j) \tag{7.30}$$

Plugging Equations (7.27), (7.28) and (7.30) into (7.26), we obtain :

$$\begin{aligned} \mathcal{D}_u(i) &= \sum_{(j,k)\in\Omega_u(i)} \delta_{jk} = \sum_{(j,k)\in\Omega_u(i)} q_j + \mu_{jk} = \\ &= \sum_{(j,k)\in\Omega_u(i)} \left\{ \left[\sum_{p\in\Gamma(j)} \alpha_j(p)\theta_j \frac{\tau}{p_r(1-\beta_j^p)} \right] + \frac{\tau}{p_r(1-\beta_j^k)} \right\} \end{aligned} \tag{7.31}$$

$$\mathcal{D}_d(i) = \sum_{(j,k)\in\Omega_d(i)} \delta_{jk} = \sum_{(j,k)\in\Omega_d(i)} q_j + \mu_{jk} = \\ = \sum_{(j,k)\in\Omega_d(i)} \left\{ \left[\sum_{p\in\Gamma(j)} \alpha_j(p)\theta_j \frac{\tau}{p_r(1-\beta_j^p)} \right] + \frac{\tau}{p_r(1-\beta_j^k)} \right\} \tag{7.32}$$

7.5 Numerical Results

Once the mathematical details of the model are explained, it is important to describe the performance results that can be obtained by means of its implementation. Figure 7.4 shows a block diagram of the performance analysis framework proposed in this paper.

First, in the initialization phase, the topology is set up and the shortest paths are computed. All the parameters introduced in the Sections 7.3-7.4 are chosen together with tolerance ϵ. We assume that all the nodes start with empty buffers at $t = 1$; for this reason, at the first time-slot it is not necessary to calculate the *beginning* state probabilities :

$$\pi_B(i, 0, 1) = 1 \qquad \forall i \in V$$
$$\pi_B(i, n, 1) = 0 \qquad \forall n \in (0, Z], \forall i \in V$$

Then, the state probabilities in the *beginning*, *transmission* and *end* phases for each time-slot are iteratively computed, as per equations (7.20)-(7.22), (7.6)-(7.8), and (7.15)-(7.17) respectively. Iterations stop when either the norm $\Delta(t)$ of the difference between two consecutive steps is lower than ϵ or the maximum number of iterations is performed (i.e., 1000 in our implementation).

If a stationary point is found, we calculate the delays, according to (7.32) and (7.31), which allows the feasibility assessment of smart grid applications.

However, if a stationary point is not found, several analysis can be performed with the presented model. One of the possible applications is to evaluate potential bottlenecks of the system, for instance nodes whose delay is constantly increasing or congestioned areas with collision probability higher than the average.

The proposed analytic model and the performance evaluation framework are applied to an AMI system in Villeray, a urban neighbourhood in Montreal, in order to show their capabilities and to describe the different types of performance analysis that can be carried out. The considered topology is composed of 2 data collectors, 16 routers and 6033 smart meters. The location of routers and collectors was found in Hydro-Québec (2012), a report of a pilot-project of smart-meter deployed in Quebec, whereas the position of the smart meters

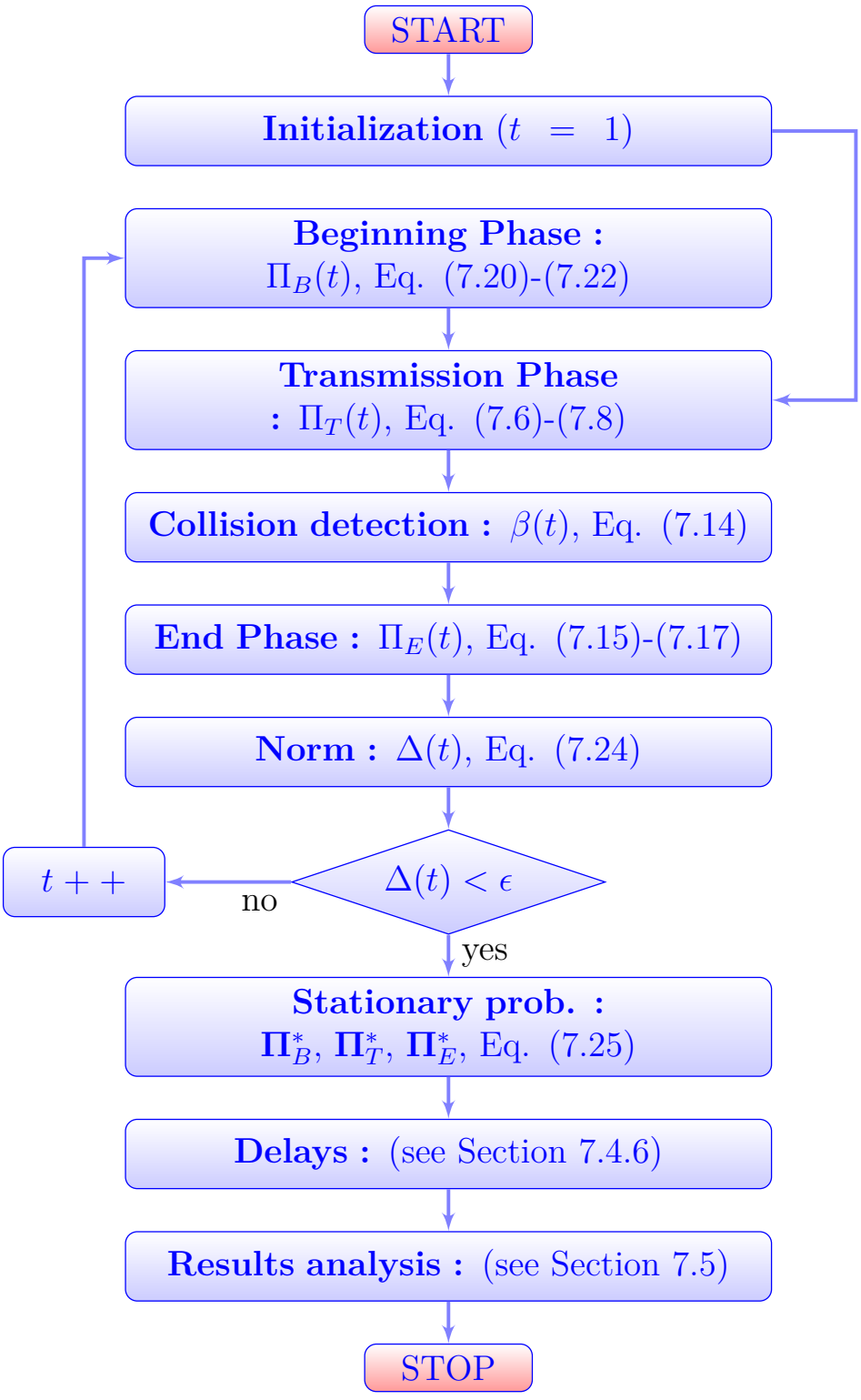


Figure 7.4 Block diagram of the performance analysis framework.

was derived from a publicly available dataset of residential addresses in the Montreal area[3]. The assumption of one smart meter per residential address was adopted. The links in the topology were defined using a covering ray of 1500 m for routers and collectors, and of 150 m for the smart meters.

Different traffic profiles and system characteristics (e.g., retransmission probability and buffer size) were adopted in the proposed test cases, to highlight the level of insight that it is possible to gain with the presented analytic model. Four types of results were extracted and will be discussed in the next sections.

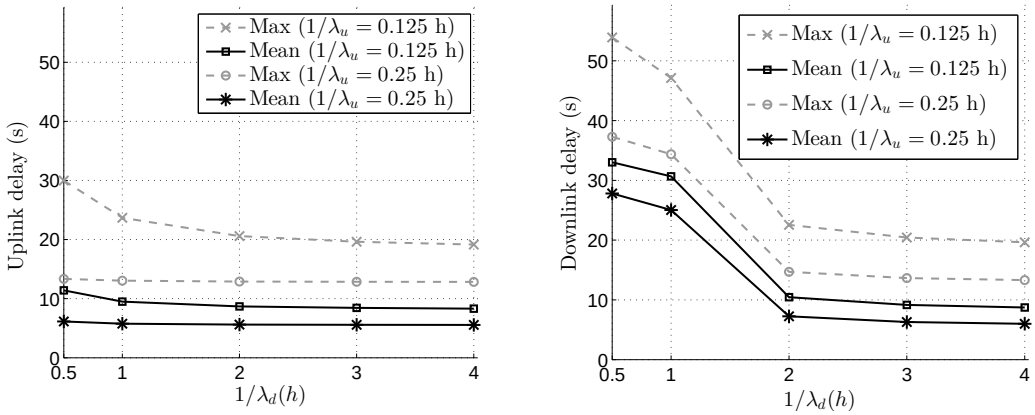
7.5.1 Feasibility assessment

One of the objectives of this research is to enable feasibility assessments for different kinds of applications in various fields, such as Smart Grids or IoT. A generic application can be characterized by the traffic it will generate within the communication network. Therefore, each application can be considered as a separate traffic commodity. The assumption of Poisson-distribution permits to aggregate the traffic of different applications by summing up their average traffic generation rates, resulting in λ_u and λ_d. These two parameters are used as inputs to our feasibility evaluation.

In the assessment of the feasibility of a new application in a given infrastructure, the average delay is one the most important parameters. Uplink and downlink delays can help determine if one communication infrastructure is well suited for a given application. The assessment is performed by comparing the average and maximum delay requirements of the application and the average and maximum delay observed in the network. In fact, our model permits to calculate the average delay at each node, and one can easily determine the largest observed delay, and compare it to the application requirements. Some statistics on the uplink and downlink delays are reported in Figures 7.5 and 7.6 with respect to the mean packet generation times $1/\lambda_d$ and $1/\lambda_u$. Some of the generated results are omitted for the sake of clarity.

In particular, in Figure 7.5 we used a buffer size of $Z = 20$ packets, a retransmission probability $p_r = 0.5$ and $1/\lambda_u = 0.125, 0.25$ h. The uplink delay (on the left) and the downlink delay (on the right) are reported according to $1/\lambda_d$ that varies between 0.5 and 4 h. The maximum delays are represented by dashed grey lines, whereas the average delays by solid black lines. It is observed that the average and maximum delays decrease as the mean packet generation times decrease. However, the slope of the reduction decreases in the absolute value as $1/\lambda_u$

[3]A file with the list of all the residential addresses in the island of Montreal can be found at `http://donnees.ville.montreal.qc.ca/dataset/adresses-ponctuelles`



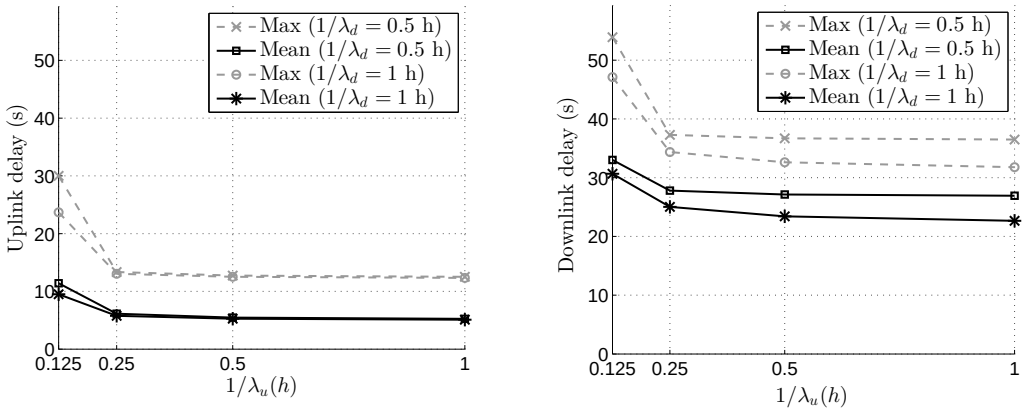
(a) uplink (b) downlink

Figure 7.5 Average and maximum of the delay with respect to the mean packet generation time in downlink ($1/\lambda_d$) with $p_r = 0.5$, $Z = 20$ packets, and $1/\lambda_u = 0.125, 0.25$ h.

and $1/\lambda_d$ decrease. In other words, when the packet generation time grows above a certain threshold, the correspondent delay reduction is lower and lower. Comparing the two figures, we also noticed that the delay is considerably larger in the downlink direction for low values of $1/\lambda_d$, whereas uplink and downlink delays are very similar for higher values of $1/\lambda_d$.

Figure 7.6 reports the average delays in uplink (on the left) and downlink (right) obtained with $Z = 20$ packets, $p_r = 0.5$ and $1/\lambda_d = 0.5, 1$ h. Delays are plotted with respect to $1/\lambda_u$, that varies between 0.125 and 1 h. The same considerations expressed for Figure 7.5 are also valid for Figure 7.6 with one exception : in Figure 7.5, the difference between uplink and downlink delays almost vanishes as $1/\lambda_d$ increases; in Figure 7.6, this difference is not affected by $1/\lambda_u$.

The reason for this difference is investigated using Figure 7.7, which presents an illustration of the shortest paths (i.e., $\Omega_u(i)$ and $\Omega_d(i)$) between the i-th smart meter and its associated collector c, through all the intermediate nodes (i.e., $x_1, \ldots, x_k$). The downlink delay $D_d(i)$ (top of the figure), the uplink delay $D_u(i)$ (bottom of the figure) and their subcomponents (i.e., the queueing delays q_j and the transmission delays μ_{jk}) are also reported. Assuming no sensitive variations in the transmission delays in the two possible directions ($\mu_{jk} \simeq \mu_{kj}$, $\forall(j,k) \in \Omega_u(i)$,$\forall(k,j) \in \Omega_d(i)$), we can infer that the difference between the uplink and downlink delays only depends on the queueing delays. Moreover, k queueing delay components (i.e., q_{x_1}, q_{x_2}, ..., q_{x_k}) are present in both downlink and uplink. The only remaining elements in this comparative analysis are the queueing delays at node i (i.e., q_i) and at its associated data collector (i.e., q_c). Since the queueing delay of a node is expected to depend on its mean packet transmission rate, and since smart meters transmit packets



(a) uplink (b) downlink

Figure 7.6 Average and maximum of the delay with respect to the mean packet generation time in uplink ($1/\lambda_u$).

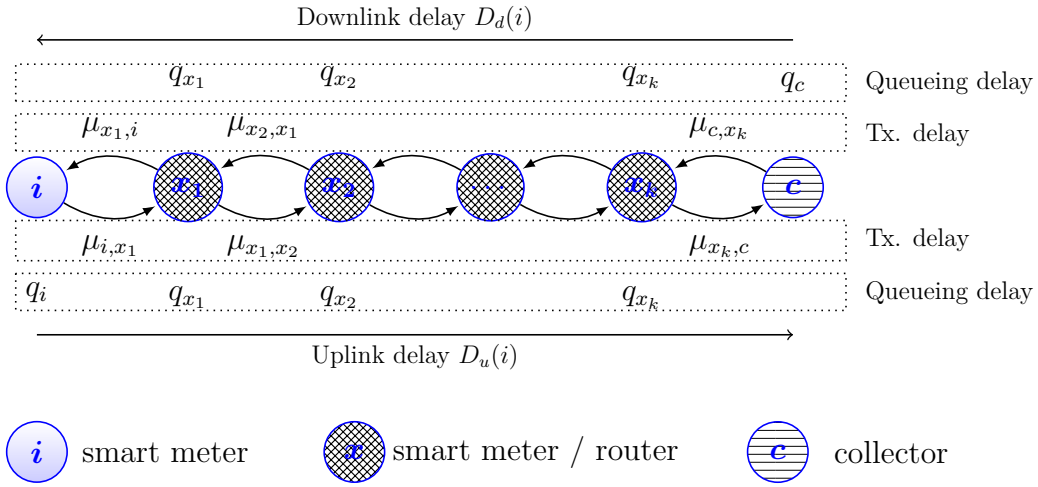


Figure 7.7 Uplink and downlink delays between smart meter i and its associated collector c.

considerably less frequently than the data collectors, we can state that the queueing delays of smart meters are substantially lower than those of data collectors (i.e., $q_i \ll q_c$). At the light of these considerations, we can state that the difference between the downlink and the uplink delay can be approximated with the queueing delay at the collector :

$$D_d(i) - D_u(i) \simeq q_c - q_i \simeq q_c \tag{7.33}$$

From Eq. (7.1), we know that the mean transmission rate of the collector increases with λ_d. As a consequence, the queueing delay increases, and according to Eq. (7.33), so does the quantity $D_d(i) - D_u(i)$. Therefore, we expect downlink and uplink delay to differ sensitively when $1/\lambda_d$ is low, and to be closer to each other when $1/\lambda_d$ is large, as confirmed in Figure 7.5. On the other hand, there is no dependence of the mean transmission time of the collectors on $1/\lambda_u$. Consequently, the difference between uplink and downlink delay is not expected to change as $1/\lambda_u$ increases or decreases, which corresponds to the observations made on Figure 7.6.

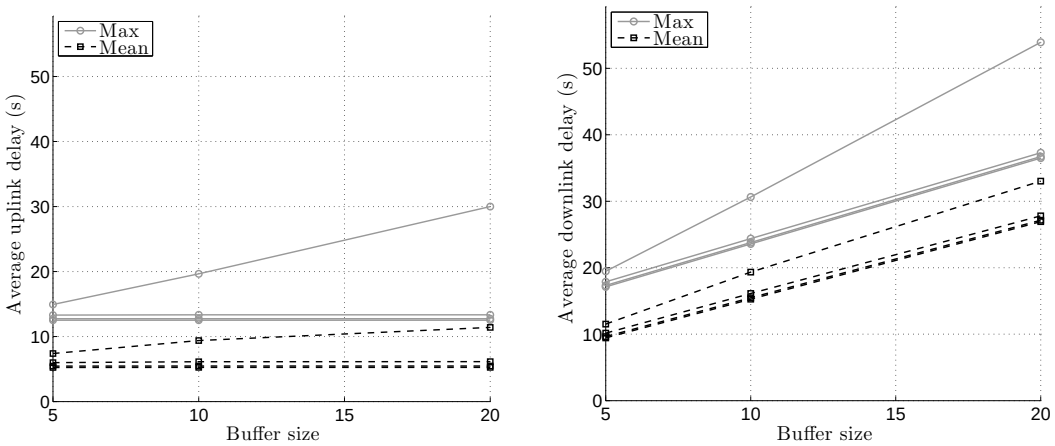
7.5.2 Impact of the buffer size

A separate analysis was carried out to highlight the impact of the buffer size Z on the uplink and downlink delays. In Figure 7.8, the average (black dashed line) and maximum (grey continuous line) values of the delay are reported with respect to 3 possible values for the buffer size Z (i.e., 5, 10, and 20 packets) and for $1/\lambda_u = 0.125, 0.25, 0.5, 1$ h (not indicated in the figure to avoid overloading). In all cases, $p_r = 0.5$ and $1/\lambda_d = 0.5$ h. As in the previous analyses, the downlink delay results to be higher than the uplink. Moreover, the downlink delay grows almost linearly as the buffer size increases, whereas the uplink delay seems to be invariant.

The use of a relatively large value for the mean packet generation time (0.5 h) implies high queueing delays at the collectors; this results in a higher expected number of packets in the buffer. Consequently, as the buffer size increases, more packets can be stored in the buffer, therefore leading to higher downlink delays. This trend is confirmed by the values of the

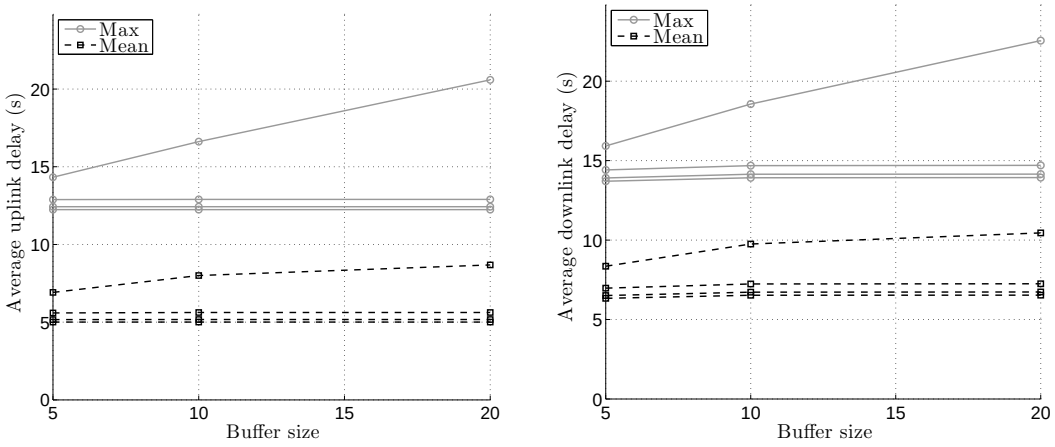
Table 7.1 Average and maximum values for the number of packets in the buffer according to different values of the buffer size Z, obtained with $p_r = 0.5$, $1/\lambda_u = 0.125$ h, and $1/\lambda_d = 0.5$ h.

Buffer occupancy (packets)	$Z = 5$	$Z = 10$	$Z = 20$
Average	0.0012	0.0017	0.0028
Maximum	2.28	5.45	11.77



(a) uplink (b) downlink

Figure 7.8 Average and maximum of the delay with respect to the buffer size with $1/\lambda_d = 0.5$ h.



(a) uplink (b) downlink

Figure 7.9 Average and maximum of the delay with respect to the buffer size with $1/\lambda_d = 2$ h.

average number of packets in the buffer for $p_r = 0.5$, $1/\lambda_u = 0.125$ h, and $1/\lambda_d = 0.5$ h, reported in Table 7.1 : an almost linear proportion is observed between the buffer size and the buffer occupancy of the collectors (see maximum values). This linear growth is reflected in Figure 7.8.

In Figure 7.9, we performed the same analysis but with an higher mean generation time in downlink ($1/\lambda_d = 2$ h). A lower congestion at the collectors is obtained, and consequently the dependence of the downlink delay on the buffer size Z is considerably lower than in the cases with $1/\lambda_d = 0.5$ h.

7.5.3 Visual bottleneck identification

The proposed performance evaluation framework also offers the possibility to carry out visual analyses. In particular, it is possible to use the geographical coordinates to report all the nodes in a heat-map. Each node is represented by a point whose color varies according to its delay. In Figure 7.10, one heat-map for the downlink (top figure) and one for the uplink (bottom figure) delays are reported. The 2 data collectors are represented with white triangles, the 16 routers with white circles, and 6033 smart meters with points in the grey-scale. A color-bar is displayed on the right side of each map to show the correspondence between the color and the delay. This heat-map was generated using $Z = 20$ packets, $p_r = 0.8$, $1/\lambda_u = 0.125$ h, and $1/\lambda_d = 0.5$ h. The first thing to remark is that the uplink delay is considerably lower than the downlink delay, as was also observed in Section 7.5.1. This depends, as previously discussed, on large queueing delays at the collectors, caused by the use of a high data transmission rate in downlink.

Moreover, while larger uplink delays are observed at nodes farther from the data collectors, downlink delays seem to be more homogeneous throughout the considered area. As was shown in Figure 7.7, the queueing delay of the collectors, which is responsible for the augmentation of the downlink delay at all the nodes, thereby justifying the observed delay homogeneity.

This type of analysis permits to visually identify potential bottlenecks in the network, and can be very useful both in the design phase of a new system and in the performance evaluation of an existing one.

7.5.4 Scalability analysis

When it comes to the performance analysis of large scale networks, it is fundamental to study how the proposed model converges, and also what happens when the size of the topology scales up. Since the model is solved iteratively, it is important to consider the norm of the

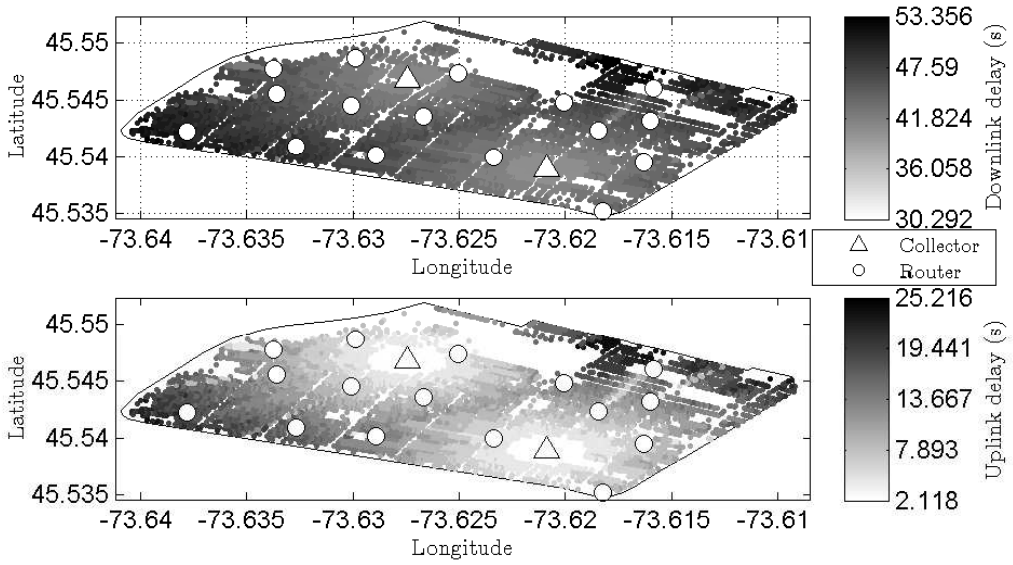


Figure 7.10 Heat-map of the delay in uplink and downlink with $1/\lambda_u = 450$ s, $1/\lambda_d = 1800$ s, buffer size of 20 packets, retransmission probability of 0.8, and 80 wireless channels.

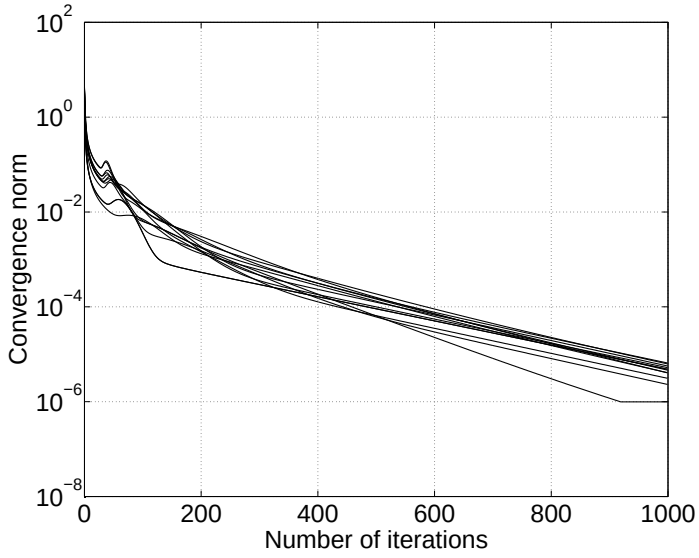


Figure 7.11 Convergence of the norm with respect to the number of iterations.

variation in the state probabilities between each pair of consecutive iterations (i.e., $\Delta(t)$ in Eq. (7.24)). The time evolution of this norm is reported in Figure 7.11 for 11 different topologies, ranging from 1715 to 11917 nodes. In all cases, the maximum number of iterations was set to 1000, the tolerance ϵ was set to 10^{-6}, and the model converged to a stationary solution with a final norm below 10^{-5}. In one case, the tolerance was reached at the 915-th iteration.

The relation between the size of the topology and the computational time was also investigated. Figure 7.12 reports the variation of the computational time according to the number of nodes (bottom figure) and of links (top figure). One can notice that the computational time tends to increase on average, as the number of nodes becomes larger. However, in few cases we observed instances with a lower number of nodes requiring a higher computational time with respect to larger topologies : e.g., the network with 3126 nodes required slightly less than 8 hours, whereas the topology with 4786 only 2 hours. In order to explain this anomaly, the computational time was also related to the number of links in the topology. A quasi-linear relation was observed between the computational time $C(i)$ of the topology i and its number of links L_i. In particular, the slope C_i/L_i is almost constant, ranging from 0.0034 to 0.0041. This implies that the number of links has a higher impact on the performance analysis computational burden, than the number of nodes. In fact, the number of links affects the wireless interference, and consequently the collision probability, whose computation proved to be intensive. However, a solution was found in less than 10 hours for 10 instances out of 11. A solution was also found for the topology with 11917 nodes, but a higher computational time was required (i.e., 18.2 hours).

7.5.5 Results validation

As discussed in Sections 7.1 and 7.2, it is difficult to find publicly available data on the RF-Mesh performance in the literature that could be used to validate our results. Therefore, we implemented a large-scale RF-Mesh network simulator, described in Malandra and Sansò (2016), and used it for validation purposes. The same topology (i.e., urban area of Villeray) was used, and we tuned the simulator parameters in order to be consistent with the results generated with the analytic model (e.g., $1/\lambda_u = 1$ h, $1/\lambda_d = 4$ h, $Z = 20$, and $p_r = 0.5$). We let the simulation time-horizon vary between 1 and 14 days and calculated the Root Mean

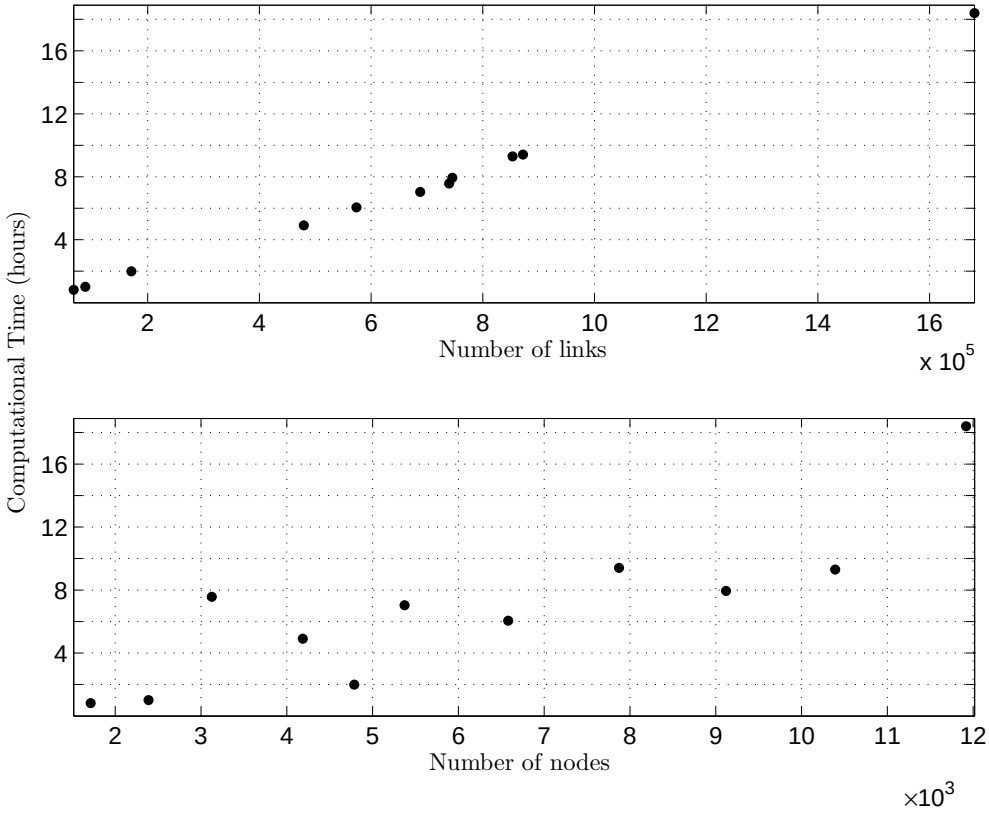


Figure 7.12 Computational time according to the number of links (top) and nodes (bottom) in the topology.

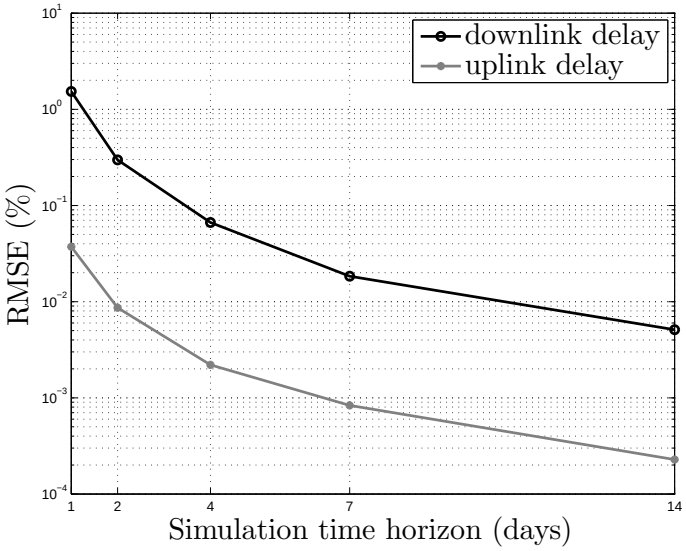


Figure 7.13 RMSE (%) between the delay computed in this work and the delay obtained with the simulator presented in Malandra and Sansò (2016) for the Villeray topology with $1/\lambda_u = 1$ h, $1/\lambda_d = 4$ h, $Z = 20$, and $p_r = 0.5$.

Square Error (RMSE) for the delay in uplink and downlink as follows :

$$RMSE_u = \sqrt{\frac{\sum\limits_{i \in N} \left(\mathcal{D}_u(i) - \mathcal{D}_u^{sim}(i)\right)^2}{|N|}} \tag{7.34}$$

$$RMSE_d = \sqrt{\frac{\sum\limits_{i \in N} \left(\mathcal{D}_d(i) - \mathcal{D}_d^{sim}(i)\right)^2}{|N|}} \tag{7.35}$$

where $\mathcal{D}_u^{sim}(i)$ and $\mathcal{D}_d^{sim}(i)$ are respectively the uplink and the downlink delay obtained with the network simulator.

The outcome of this analysis is reported in Figure 7.13. The values of $RMSE_u$ and $RMSE_d$ are reported as percentage, computed by dividing them by the mean uplink and delay values, respectively 4.975 s and 5.432 s. As one can notice, the RMSE in the downlink delay is high (i.e., 153.032 %) when the time-horizon used in the simulation is just 1 day. When we use larger time horizons, the RMSE sensitively decreases. In particular, the RMSE observed between the MMS model and a 14-day simulation was of 0.028 s in downlink and of 0.001 s in uplink, that corresponds to a relative error of 0.511% and of 0.023% with respect to the average downlink and uplink delays obtained with the simulator. This proves agreement between the analytic solution and a simulation performed on a sufficiently large time horizon.

One can also notice that the RMSE in the uplink is considerably lower than in the downlink, even with 1-day simulation horizon. This behaviour is probably due to the fact that, in the simulator, the delay is computed averaging the observed delays of all the transmitted packets; since a larger number of packets is transmitted in the uplink than in the downlink ($\lambda_u = 4\lambda_d$), a higher accuracy (lower RMSE) is found in the uplink.

7.6 Conclusions and future work

In the context of smart grid development, performance analysis is important to assess the implementability of proposed applications. In this paper, we presented a Markov-modulated system for the performance analysis of large scale RF-Mesh systems. The model takes into account relevant implementation details, such as the retransmission probability and the buffer size, which were not included in previous analytic models for the performance analysis of RF-Mesh systems. Various traffic scenarios were considered in the analysis in order to evaluate the feasibility of a wider range of possible applications. Every applications can be identified by its requirements in terms of communications (e.g., a desired average delay on the packets reception) and its feasibility can be assessed using the proposed model, as shown in the results. The use of Poisson-distributed traffic permits to carry out combined analysis of different applications on the same communication infrastructure, which is particularly sought in the Smart Grid/IoT domain, where the number of applications is large. Wireless interference was taken into account in the computation of the collision probability. An analytic formulation of the delay was also proposed.

A framework for the performance evaluation of large-scale RF-Mesh systems was established. It is based on the computation of the delay and provides meaningful insights on the performance of RF-Mesh AMIs. In particular, it permits to estimate the average delay of all the nodes in the topology with respect to the traffic generation time, and to the buffer size. A topology derived from real geographical data was used as input of the framework, to show its capabilities. Despite the large size of the instances (in the order of thousands of nodes and hundreds of thousands of links), a stationary solution was found with high accuracy (an absolute error of 10^{-5} over hundred of thousands of state variable values) with a computational time ranging from 1 to 18 hours. The use of geographical data also enables a visual representation of the delay; this feature can be used in order to identify potential bottlenecks in the design phase of a new system, resulting in consistent performance improvements.

CHAPTER 8 GEO-BASED TOPOLOGY GENERATOR

8.1 Introduction

Topology generation takes on great importance when it comes to studying the performance of AMI systems : the characteristics of the topology considerably affect the performance analysis. For instance, densely populated areas are subject to a high level of interference that needs to be properly addressed. On the other hand, the performance in less crowded countryside regions is undermined by low network connectivity and, as a consequence, network planning and design becomes fundamental to guaranteeing a suitable system operation.

Nevertheless, a good deal of research uses randomly distributed topologies, since they are simple to generate and can be easily integrated in a mathematical formulation. However, random topologies do not reflect the fact that smart meters, the principal involved nodes, are far from being randomly distributed since they are mainly installed in the residential dwellings. To this purpose, we decided to build a dedicated module to get more consistent topologies.

Another important issue with AMI topology generation is size : deployed AMIs are composed of several thousands of nodes. Moreover, the position of nodes, other than smart meters, are not easily retrieved in publicly available maps, and in some cases (i.e., for routers and collectors) they are undisclosed and protected by confidentiality agreements.

The topology generation is composed of two phases : (i) the node placement, and (ii) the link definition, which are respectively described in Sections 8.2 and 8.3.

8.2 Nodes placement

The topology generation tool is used to find the location of the set of nodes V, which includes the set M of smart meters, the set R of routers, and C of data collectors. The three following options are implemented :

- use of existing data
- manual draw
- use of open-source databases

In the first case, the user provides a file with the list of node locations. Then, the points are represented and can be visualized on a map. The second possibility is to let the user manually draw the position of nodes in a map. Then, the latitude and the longitude of the selected points are retrieved using Google Maps APIs. The manual choice of the points might be a good option for routers and collectors, which are in the order of few tens. On the other hand, smart meters, which can be in the thousands for large AMI implementations, are far too numerous to be manually selected on a map. As a consequence, the third option is more suitable to locate smart meters, based on the use of publicly available information (the third use mode) about residential addresses (e.g., list of postal codes, municipality's databases) to create the topology.

The different options can be combined in order to provide the user with more flexibility in the choice of the topology : e.g., an instance can be generated using the smart meter location of a previously generated topology and manually drawing routers and collectors on it. In Figure 8.1, an example of the creation of a topology is described in more detail. We chose to focus our analysis on a neighborhood in the municipality of Montreal, for which we found a database of all the residential addresses[1] : the user can draw a polygon to select an area (see Figure 8.1(a)), then a script permits to compute the set of smart meters within the selected area; then, the user is then asked to draw on the map the position of routers (see Figure 8.1(b)) and data collectors (see Figure 8.1(c)).

Each of the three options permits to generate a file with the list of nodes and their GPS coordinates. The file, in which the lines are structured as "ID, node type, latitude, longitude", is called "Topology.txt" and is used in the link definition.

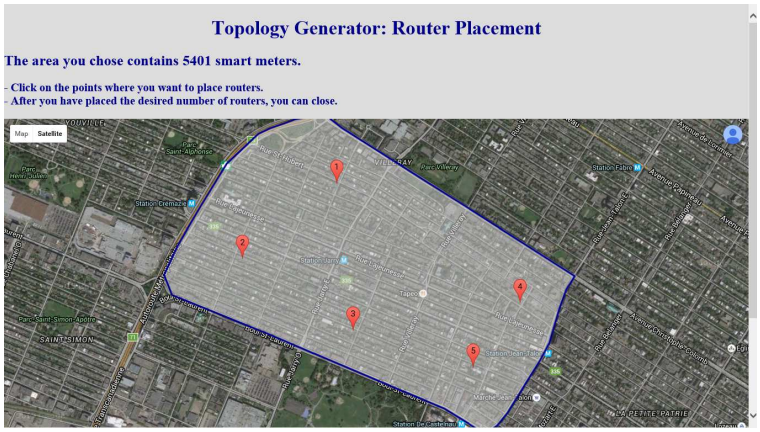
8.3 Link definition

The link definition phase completes the topology creation process. The location of each node i in the topology can be retrieved by the file "topology.txt" in the form $lat(i), long(i)$ and is used to compute the distance $d(i,j)$ between each pair of nodes (i,j), according to the so-called Haversine formula (Robusto (1957)) :

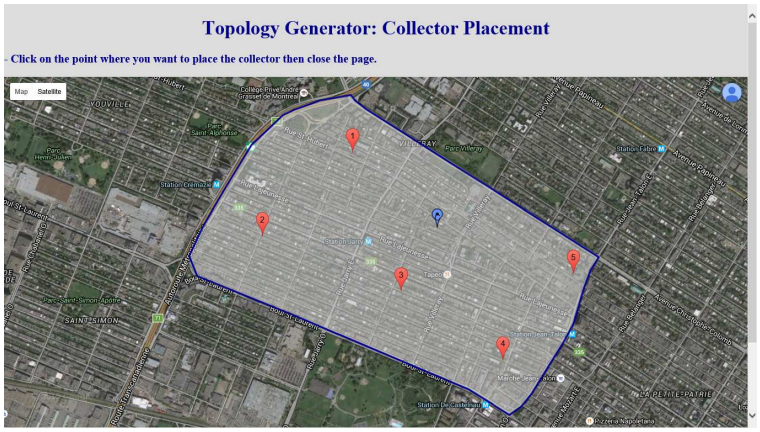
$$d(i,j) = 2r \arcsin\left(\sqrt{\sin^2\left(\frac{\phi_j - \phi_i}{2}\right) + cos(\phi_i)cos(\phi_j)\sin^2\left(\frac{\psi_j - \psi_i}{2}\right)}\right) \tag{8.1}$$

[1] A file with the list of all the residential addresses in the island of Montreal can be found at `http://donnees.ville.montreal.qc.ca/dataset/adresses-ponctuelles`

(a)



(b)



(c)

Figure 8.1 Illustration of the topology creation with the definition of a new area (Figure 8.1(a)), the choice of routers (Figure 8.1(b)), and collectors (Figure 8.1(c)).

where ϕ_i, ϕ_j are the latitudes (i.e., $lat(i)$ and $lat(j)$ respectively) expressed in radians, and ψ_i, ψ_j are the longitudes (i.e., $lng(i)$ and $lng(j)$ respectively) expressed in radians.

The set of links E is then defined using the two covering rays R_m, and R_r : the first for the smart meters, the second for routers and collector. As shown in Figure 8.2, R_r is usually larger than R_m because routers and data collectors are equipped with better radios than smart meters and experience better propagating conditions, since they are installed in privileged positions (e.g., on the top of a building). Numerical values for the two covering rays vary according to the chosen topology (e.g., larger values in a rural environment, and shorter in crowded urban topologies).

Covering rays are then used to determine, as showed in Figure 8.3, whether the link $e_{ij} \in E$ exists or not according to the following rule, based on the distance $d(i, j)$ and on the covering rays of the two nodes $R(i)$ and $R(j)$:

$$E = \{e_{ij} : d(i,j) \leq \min\left(R(i), R(j)\right) \quad \forall i, j \in V\} \tag{8.2}$$

This rule implies that a link between two nodes exists if and only if each of them is in the covering area of the other node. The link definition process yields a text file (i.e., *links.txt*) with the list of all the links considered in the topology. The list is stored in the form $i; j; d(i, j)$.

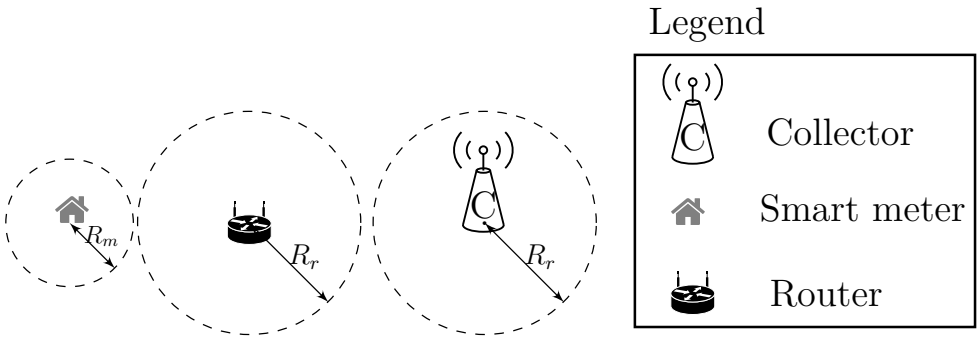


Figure 8.2 Covering areas of smart meters, routers, and data collectors.

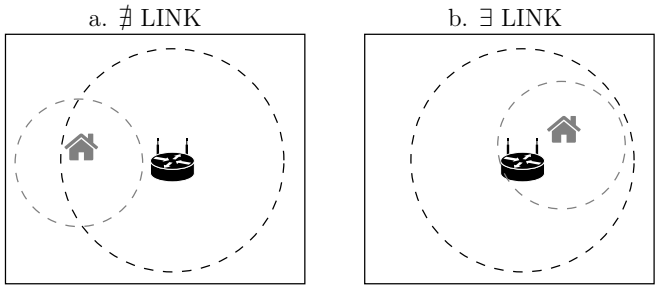


Figure 8.3 Examples of link definition between a router and a smart meter.

CHAPTER 9 SMART GRID CASE STUDY : THE SmartDESC PROJECT

The performance evaluation framework object of this thesis was developed in the context of the SmartDESC project, a load-management project partially funded by Natural Resources Canada, and carried out by a team of professors and students in different areas, such as mechanical engineering, power system, control, linear and non-linear optimization, and telecommunications.

The two main objectives of the project were (i) to shave the peak of power demand and (ii) to ease the integration of energy coming from renewable sources. This was achieved by using the large population of EWHs in Québec (i.e., nearly 4 millions) as a huge battery to be used to store energy in case of excess power production (especially the wind power), and also to reduce the consumption in peak hours by shifting in time the water-heating process.

A simplified SmartDESC architecture, including only the relevant components for the telecommunications module, is reported in Figure 9.1. However, a detailed report of the project, describing the general structure, the modelling details of the different research area, and the achieved results can be found in Sirois et al. (2016). As one can notice, each EWH is equipped with a mean-field-based local controller (Kizilkale and Malhame (2014)). A central entity, called "scheduler", is in charge of determining the target temperature, which is the information to be transmitted to each EWH (Tammam (2016)). Moreover, at randomly distributed time instants, each EWH sends back to the scheduler information on its current temperature and state of charge (Solis (2015)).

As shown in Figure 9.1, the bidirectional communication between the scheduler and the EWHs is achieved by a central telecommunications module, represented by a blue-filled rectangle. The resulting star topology (i.e., with the telecommunications node directly connected to all the other nodes in the architecture) represents the logical telecommunications architecture of the SmartDESC project. However, the telecommunications system actually implemented is not characterized by a star topology, because it would be unfeasible to have a direct connection between the scheduler and the large multitude of EWHs, distributed over a very extended and heterogeneous geographic area (i.e., the province of Québec). Details about the AMI actually implemented in Québec can be found in Hydro-Québec (2012), a report of a pilot project installation of smart meters in Québec, and in Lichtensteiger et al. (2010), where the RF-Mesh system AMI under study is thoroughly described.

The main purpose of the telecommunications part of the SmartDESC project was to study the performance of the telecommunications system in order to verify whether the data transmis-

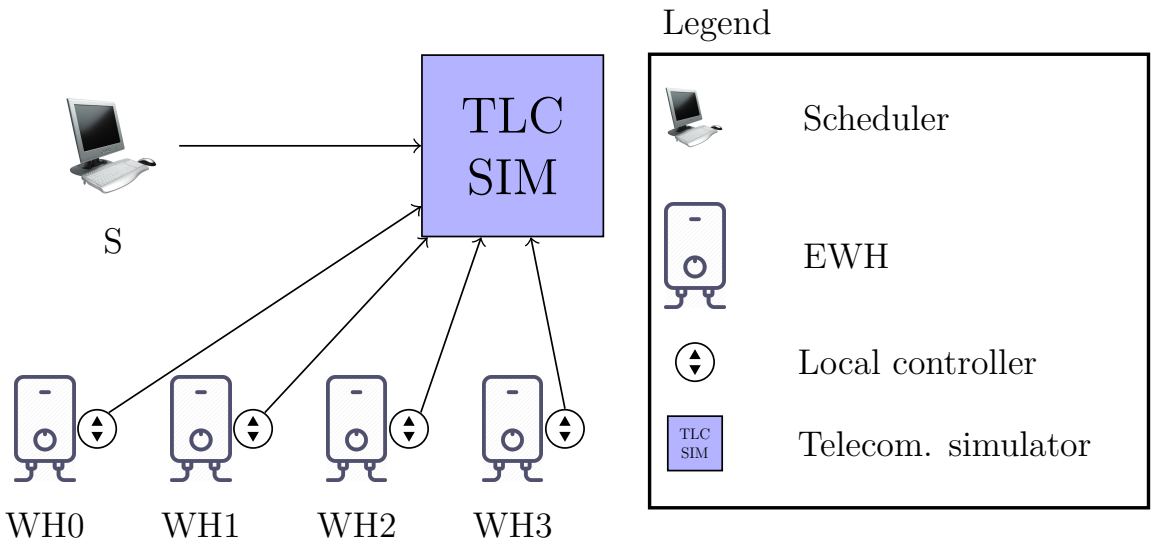


Figure 9.1 Simplified architecture of the SmartDESC project with the scheduler S, 4 EWHs, and the telecommunications module (TLC-SIM).

sions required by the SmartDESC project could be supported by the existing communication infrastructure. Packet delay was chosen as the basic performance evaluation index . In Figure 9.2, an illustration of the end-to-end delay between an EWH and the scheduler is reported. The delay can be decomposed in three sub-components :

- HAN, the delay from the EWH to the smart meter
- NAN, the delay from the smart meter to the data collector
- WAN, the delay from the data collector to the Scheduler

The first and the third layers of the architecture use standard technologies and protocols, whose performance is well studied in the literature : in particular, Zigbee prevails in HAN communications, and satellite and cellular technologies are predominant in WAN communications. Both have a small (i.e., in the order of hundreds of milliseconds) and stable delay. On the other hand, as discussed in Chapter 4, 5, 7, and 6, the delay in RF-Mesh NAN is considerably higher and can take values in a very large range : it was observed in Hydro-Québec (2012) that the average delay varies between 1 s and 18 min, 43 s. As a consequence, the NAN component of the delay is the object of our study.

In order to evaluate the network delay in the context of the SmartDESC project, a slightly modified version of the simulator presented in Chapter 6 was used. Figure 9.3 is helpful to

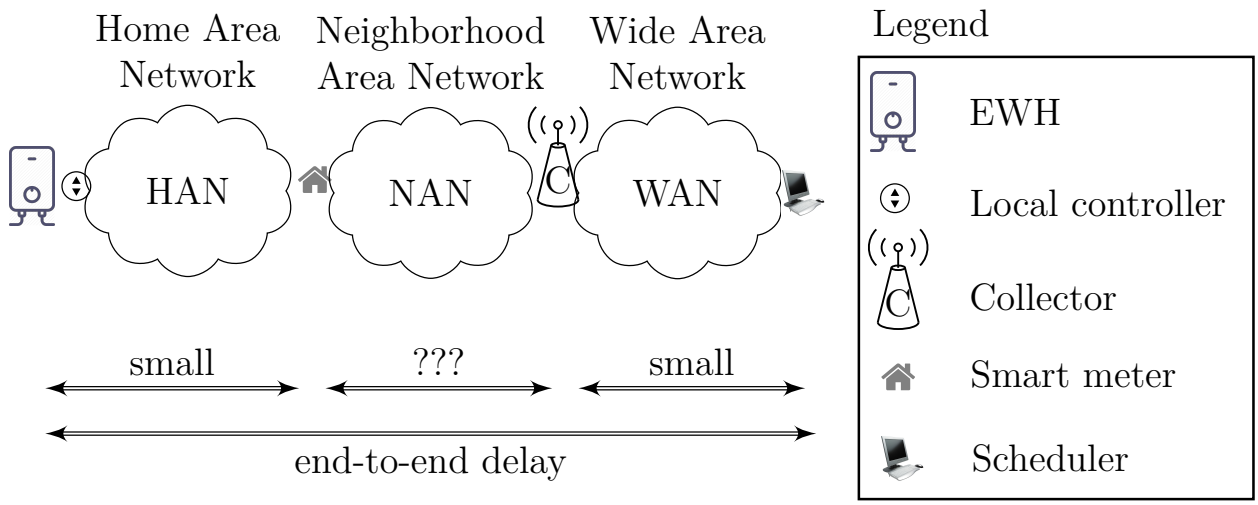


Figure 9.2 Representation of the end-to-end delay between a EWH and the scheduler, in the SmartDESC architecture.

clarify the link between the logical SmartDESC architecture, akin that illustrated in Figure 9.1, and the implemented network simulator. The figure is composed of two boxes : the top one contains the logical structure of a simplified SmartDESC architecture with 5 EWHs (i.e., $WH0, \ldots, WH4$, a scheduler S, and the TLC-SIM module; the bottom box contains the associated telecommunications structure, with 5 SmartDESC-controlled smart meters (gray houses), 2 regular smart meters (black houses), one router, and one data collector (connected to the scheduler). Dotted arrows connect corresponding elements in the SmartDESC architecture and in the network simulators. As one can notice, the router has no corresponding element in the SmartDESC architecture, since it is a characteristic device of the communication system which has no other functions than packet relaying. The difference between SmartDESC-controlled — that correspond to an EWH — and regular smart meters will be explained below.

Two major issues came up during the integration of the network simulator into the SmartDESC simulator : first, the size of the instances (i.e., the number of EWHs) used in the SmartDESC project is limited to several hundreds, which is considerably too small to be representative for a RF-Mesh NAN; second, the SmartDESC simulator uses time-steps of 5 min (hereafter referred to as *big step*), whereas the network simulator works with time steps of 0.7 s (hereafter referred to as *short step*). The difference in the simulation structures of SmartDESC and AMI is due to the characteristics of these two systems. On the one hand, simulation and optimization of the different components of the SmartDESC project with more than a

few hundred nodes or with a time step shorter than 5 min was found to be prohibitively expensive. On the other hand, this structure cannot represent the behavior of RF-Mesh AMIs : first, they are composed of thousands of smart meters, and the characteristic interference behavior cannot be captured with just a few hundred nodes; second, the smart meters inherently operate in a synchronized manner each 0.7 s (as explained in the treatment of FHSS). In order not to alter the nature of the two simulators, it was decided to retain the two different simulation steps and to adopt two different types of EWHs, so that the SmartDESC simulator can work with a few hundred nodes, while the network simulator takes into account a larger population.

The two different types of EWHs considered in the network simulation are (i) SmartDESC-controlled nodes and (ii) regular nodes. Their difference lays in the traffic generation : SmartDESC-controlled nodes do not produce packets unless they are received by the corresponding nodes in the SmartDESC simulation, whereas regular nodes operate as if no SmartDESC simulation was put in place. At the end, the average delay is computed for SmartDESC-controlled nodes in order to verify that packets are received in a timely manner, but it is also computed for the regular nodes in order to check if the SmartDESC traffic has altered the regular operation of the underlying communication system.

In Figure 9.4, a flowchart of the network simulator is reported : the blue blocks (on the left side of the flowchart) contain network simulator operations, while the yellow ones (on the right side of the flowchart) represent the interactions with the SmartDESC simulator. At the beginning, the network simulator is initialized, and the variable t_{sim}, which represents the counter of *big steps* in the simulation, is set to 0. Then, the network simulation is carried out, synchronously with the SmartDESC simulator. The block "SIMULATE TLC" indicates that the network simulation is performed for a *big step*, amounting to nearly 429 *short steps*. Before this operation, the SmartDESC simulator produces a list of packets to be transmitted during the next 5 min. The list is passed to the network simulator, which simulates the operation of a RF-Mesh AMI for 429 *short steps* and produces a list of all the packets that are correctly received at their destination in this time-frame; for all the successfully received packets, the arrival time is stored and the delay is computed. This list is passed to the SmartDESC simulator, which can use the information contained in the packets (e.g., the target temperature, the temperature and the state of charge of EWHs) for its dynamics, and the information on the delay to verify if the data are still up-to-date.

Simulations were performed on a 7-day time horizon with instances of 3300 nodes, 300 of which were SmartDESC-controlled. At the beginning of each day, the scheduler sent the target temperature to each EWH in the topology (both SmartDESC-controlled and regu-

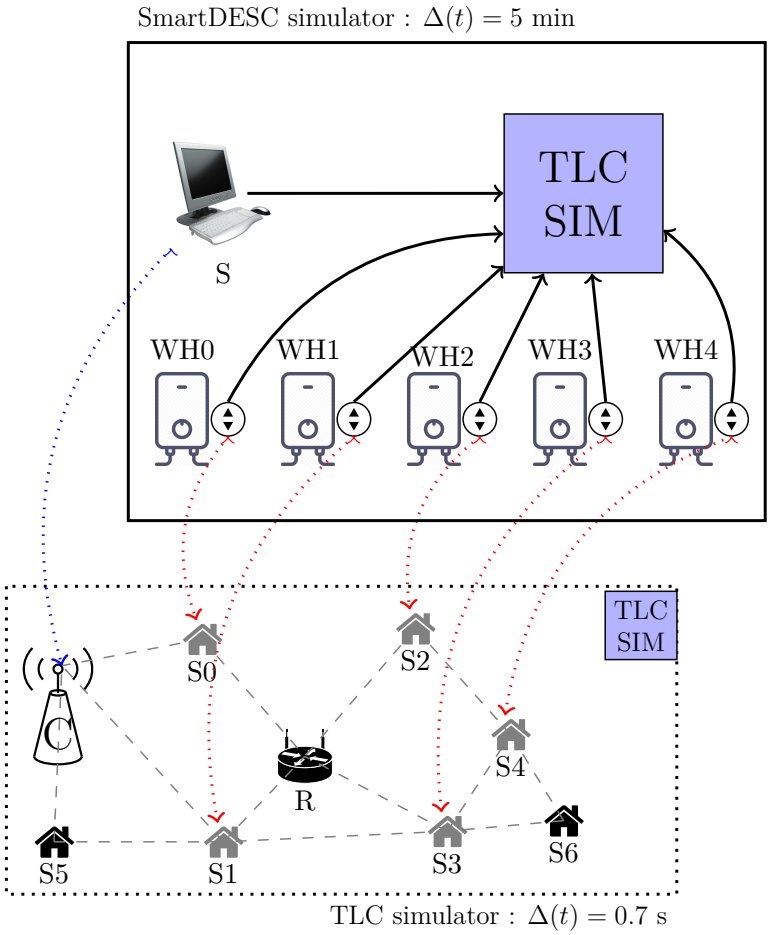


Figure 9.3 Illustration with a simple SmartDESC architecture (top box, with 5 EWH, the scheduler and the TLC simulation module), and the associated architecture of the telecommunications (bottom box, with 5 SmartDESC-controlled smart meters, 2 regular smart meters, 1 data collector, and 1 router).

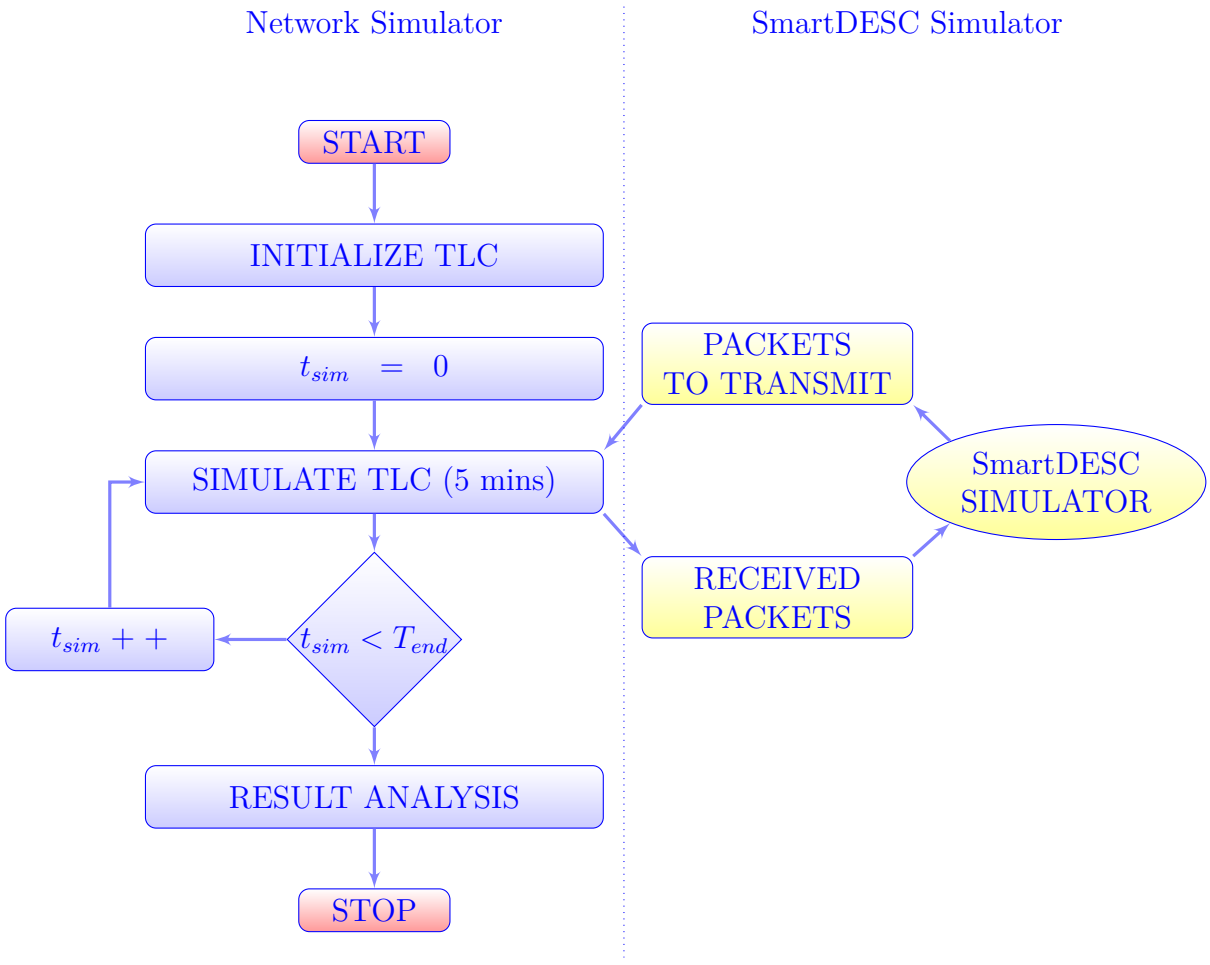


Figure 9.4 Flowchart of the network simulator for the SmartDESC application.

lar). On the other hand, each EWH sent back to the scheduler its current temperature and state of charge, according to a Poisson-distribution with an average of 4 packets per day per EWH. Broadcast transmission is not currently possible on the RF-Mesh under study, and was replaced by unicast transmissions from the scheduler to each smart meter; this creates congestion at the scheduler side, due to the quasi-simultaneous transmission of a large number of packets. As a consequence, it was observed that the broadcast downlink delay was considerably higher than the uplink delay, as shown in Table 9.1. The larger delays in the downlink directions are also due to the fact that uplink transmissions are spread over time, since they are randomly generated by each local controller; instead, downlink packets are generated at the same time for all the nodes. Finally, the average and maximum activity time of each device showed a low bandwidth utilization for the SmartDESC project. A maximum of 22% was observed at the data collector, while smart meters were transmitting for less than 0.1% of the time, on average. Overall, the introduction of SmartDESC traffic did not seem to produce any remarkable degradation of the performance with respect to the regular operation of RF-Mesh AMI, as can be observed by comparing the results in 9.1 and in Chapter 6.

Table 9.1 Table with numerical results for the telecommunications in SmartDESC simulations, with 400 SD-controlled nodes in a network of 3300 nodes, and with one *broadcast* transmission and 4 uplink transmissions on average per day.

Average uplink delay	19 s
Maximum uplink delay	38 s
Average downlink delay	303 s
Maximum downlink delay	414 s
Average activity time	0.13%
Maximum activity time	22%

CHAPTER 10 GENERAL DISCUSSION

10.1 Synthesis of the work

AMI is considered a key element of the current and future Smart Grid applications. A wide variety of technologies are used to support AMI, and RF-Mesh seems to be very popular in this context. Despite their popularity in actual AMI implementations, the performance of RF-Mesh has not been adequately studied and defined in the literature. The lack of performance evaluations is a roadblock to the adoption of RF-Mesh systems for applications other than remote reading, that was the original purpose of this kind of network. In this thesis we propose a framework for the performance evaluation of large scale RF-Mesh, taking into account aspects that were neglected by previous literature. Two methodologies (i.e., mathematical analysis and network simulations) were followed to model RF-Mesh systems and to represent their peculiarity in an accurate fashion. The two approaches have different characteristics and might result complementary in the sense that the limits of one can be overcome by the advantages of the other, and vice versa.

The main contributions of this thesis to the knowledge in the RF-Mesh performance literature are :

Geo-based topology generator

A topology generation tool, presented in Chapter 8, was implemented to characterize the input instances for the performance evaluation framework. This tool produces topologies that accurately reproduce the AMI environment, since it is based on real geographic data. The topology generator gives to the user freedom in its analysis, since it provides several use modes to accommodate a wide range of requests : in particular, the user can (i) exploit GIS data to create new instances, (ii) use previously created topologies, and (iii) manually choose node locations to test new design solutions.

Two analytic model for mesh networks with s-ALOHA and FHSS

Two analytic models, described in Chapter 4 and 5, were proposed to start the performance evaluation of large-scale RF-Mesh networks. The system under analysis is characterized by a large number of nodes, a shortest path routing, a s-ALOHA base MAC layer and the FHSS protocol. To the best of our knowledge, no other analytic study was found in literature simultaneously considering all the aforementioned aspects. A novel analytic formula

to compute the packet collision probability was derived : the formula was then used to estimate the average end-to-end delay in the communications from and to all the smart meters in the network. Performance evaluations of large-scale instances were carried out, highlighting the importance of the FHSS protocol, without which RF-Mesh AMIs would suffer a high number of collisions, and large delays. Different traffic loads were used in the performance evaluations, allowing the feasibility assessment of different Smart Grid applications. Visual analyses are also enabled by the use of GIS data.

MMS performance model considering retransmission probability and buffer size

The third analytic model, described in Chapter 7, was formulated in order to include in the modelling some important aspects (e.g., retransmission probability and buffer size), that were not taken into consideration by previous RF-Mesh AMI performance analysis. This model also provides a more accurate representation of the collision probability, which is achieved by not simply considering the transmission probability but also the destination of the transmission. Despite the large number of variables, the model was observed to converge to a stationary solution in reasonable times.

Network simulator

A network simulator was developed from scratch using well known open-source programming languages, such as Java and Python. The tool permits to validate the analytic result, and to include in the modelling complex features, which are difficult to represent in a mathematical model : in particular, a non-static routing protocol (i.e., the *layer-based* routing) was introduced, and other types of traffic distribution were considered (e.g., broadcast). The tool proved to be computationally efficient and capable of simulating large-scale instances, faster than the mathematical model.

Application to a Smart Grid project

A modified version of the network simulator was developed and used to evaluate the performance of a RF-Mesh AMI in the SmartDESC project. It allowed to validate the proof of concept proposed within the project from a communications perspective. The main results were : (i) the communication system under study was able to support the SmartDESC traffic, (ii) SmartDESC traffic did not cause a degradation in the performance of the communication system.

CHAPTER 11 CONCLUSION AND RECOMMENDATIONS

Even though it brought some important contributions to the RF-Mesh performance literature (thoroughly highlighted in Chapter 10), this thesis is subject to some limitations that need to be addressed by future studies in order to improve the quality of the research.

First of all, the routing mechanism in the analytic models is static and based on the shortest paths computation : the use of more complex routing protocols might improve the capacity to catch the dynamic behavior of mesh networks.

Second, the link definition is based on the use of covering rays : this does not allow to analyze peculiar characteristics of the propagation in the wireless environment. It would be interesting to create topologies that dynamically change (e.g., with link or node failures), or to define links according to more complex propagation models, that might explicitly evaluate the propagating conditions in a geographic area (e.g., presence of buildings and other obstacles to the propagation). Moreover, the use of field trials could permit to validate of the proposed analytic and simulation models, which have been only compared to each other in Section 7.5.5.

Additionally, the use of GIS data in the topology generator is subject to the availability of a database with the location of smart meters, which might be difficult to retrieve for some geographic areas.

Further, in the end-to-end delay representation, the HAN and WAN components were not considered. It is acknowledged that the delay in the NAN was observed to be higher and, most importantly, with a higher range of variation, but formally considering the two other components of the delay would complete the delay analysis and add value to the results.

Another possible improvement could come from the integration of several NANs in the same analysis. This would permit to evaluate the performance of the WAN, which could experience congestion due to the increasing number of new applications.

Moreover, the research followed two approaches (i.e., mathematical analysis and simulation) but it was not put in place any field-trial, which could have enriched the performance evaluation framework and enable a wide range of new analyses.

Finally, this framework includes useful tools for the performance analysis but does not consider the design phase. The location of routers and data collectors is assumed to be fixed and known in advance. A considerable improvement, from a power utility's point of view, might be achieved by integrating an optimal design model for the location of routers and

collectors which, for instance, could be able to answer to fundamental questions, such as : (i) how many routers/collectors are needed to cover a certain area? (ii) Where should they be located in order to expect a delay lower than a certain threshold?

BIBLIOGRAPHY

F. Aalamifar, L. Lampe, S. Bavarian, and E. Crozier, "WiMAX technology in smart distribution networks: Architecture, modeling, and applications," in *2014 IEEE PES T&D Conference and Exposition*. Institute of Electrical & Electronics Engineers (IEEE), Apr 2014. [Online]. Available: http://dx.doi.org/10.1109/TDC.2014.6863432

G. Abdulla, "The deployment of advanced metering infrastructure," in *Smart Grid and Renewable Energy (SGRE), 2015 First Workshop on*, March 2015, pp. 1–3. [Online]. Available: http://dx.doi.org/10.1109/SGRE.2015.7208738

N. Abramson, "The ALOHA system: another alternative for computer communications," in *Proceedings of the November 17-19, 1970, fall joint computer conference*. ACM, 1970, pp. 281–285. [Online]. Available: http://dx.doi.org/10.1145/1478462.1478502

K. Akkaya, K. Rabieh, M. Mahmoud, and S. Tonyali, "Customized certificate revocation lists for IEEE 802.11s-based smart grid AMI networks," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2366–2374, Sept 2015. [Online]. Available: http://dx.doi.org/10.1109/TSG.2015.2390131

S. Ali, M. Jawad, B. Khan, C. Mehmood, N. Zeb, A. Tanoli, U. Farid, J. Glower, and S. Khan, "Wide area smart grid architectural model and control: A survey," *Renewable and Sustainable Energy Reviews*, vol. 64, pp. 311 – 328, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032116302209

Z. Alliance, "About Z-Wave," 2011. [Online]. Available: http://www.z-wave.com

A. AlMajali, A. Viswanathan, and C. Neuman, "Analyzing resiliency of the smart grid communication architectures under cyber attack," in *5th Workshop on Cyber Security Experimentation and Test*, 2012. [Online]. Available: https://www.usenix.org/system/files/conference/cset12/cset12-final13.pdf

P. Amaro, R. Cortesao, J. Landeck, and P. Santos, "Implementing an advanced meter reading infrastructure using a Z-Wave compliant wireless sensor network," in *Energetics (IYCE), Proceedings of the 2011 3rd International Youth Conference on*, July 2011, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/6028316/

E. Ancillotti, R. Bruno, and M. Conti, "The role of the RPL routing protocol for smart grid communications," *IEEE Commun. Mag.*, vol. 51, no. 1, pp. 75–83, jan 2013. [Online]. Available: http://dx.doi.org/10.1109/mcom.2013.6400442

A. M. Andronov and V. M. Vishnevsky, "Markov-modulated continuous time finite markov chain as the model of hybrid wireless communication channels operation," *Automatic Control and Computer Sciences*, vol. 50, no. 3, pp. 125–132, 2016. [Online]. Available: http://dx.doi.org/10.3103/S0146411616030020

A. Anzalchi and A. Sarwat, "A survey on security assessment of metering infrastructure in smart grid systems," in *SoutheastCon 2015*, April 2015, pp. 1–4. [Online]. Available: http://dx.doi.org/10.1109/SECON.2015.7132989

D. Arias and G. Rodriguez, "Performance of advanced metering infrastructure using cellular communication based on uplink CDMA," in *Intelligent Computing and Internet of Things (ICIT), 2014 International Conference on*, Jan 2015, pp. 111–116. [Online]. Available: http://dx.doi.org/10.1109/ICAIOT.2015.7111550

A. Aruzuaga, I. Berganza, A. Sendin, M. Sharma, and B. Varadarajan, "Prime interoperability tests and results from field," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, Oct 2010, pp. 126–130. [Online]. Available: http://dx.doi.org/10.1109/SMARTGRID.2010.5622029

G. Athanasios and P. G. Cottis, "Exploiting the cellular infrastructure for data transmission in smart metering systems," *Journal of Computer and Communications*, vol. 4, no. 07, p. 31, 2016. [Online]. Available: http://dx.doi.org/10.4236/jcc.2016.47005

D. Ayala and G. Rodriguez, "Cellular network for uplink AMI. state-of-art CDMA and path-loss models," in *ANDESCON, 2014 IEEE*. IEEE, 2014, pp. 1–1. [Online]. Available: http://dx.doi.org/10.1109/ANDESCON.2014.7098566

R. Bayindir, I. Colak, G. Fulli, and K. Demirtas, "Smart grid technologies and applications," *Renewable and Sustainable Energy Reviews*, vol. 66, pp. 499–516, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.rser.2016.08.002

R. K. Bhatia and V. Bodade, "Defining the framework for wireless-AMI security in smart grid," in *Green Computing Communication and Electrical Engineering (ICGCCEE), 2014 International Conference on*, March 2014, pp. 1–5. [Online]. Available: http://dx.doi.org/10.1109/ICGCCEE.2014.6921383

D. Bian, M. Kuzlu, M. Pipattanasomporn, and S. Rahman, "Analysis of communication schemes for advanced metering infrastructure (AMI)," in *2014 IEEE PES General Meeting | Conference & Exposition*. Institute of Electrical & Electronics Engineers (IEEE), jul 2014. [Online]. Available: http://dx.doi.org/10.1109/PESGM.2014.6939562

S. Bluetooth, "Bluetooth core specification version 4.0," 2010, specification of the Bluetooth System. [Online]. Available: https://www.bluetooth.org/docman/handlers/downloaddoc.ashx?doc_id=229737

A. Carleial and M. Hellman, "Bistable behavior of ALOHA-type systems," *Communications, IEEE Transactions on*, vol. 23, no. 4, pp. 401–410, 1975. [Online]. Available: http://dx.doi.org/10.1109/TCOM.1975.1092823

M. Carlesso, A. Antonopoulos, F. Granelli, and C. Verikoukis, "Uplink scheduling for smart metering and real-time traffic coexistence in LTE networks," in *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015, pp. 820–825. [Online]. Available: http://dx.doi.org/10.1109/ICC.2015.7248423

S. Cespedes, A. Cardenas, and T. Iwao, "Comparison of data forwarding mechanisms for AMI networks," in *Innovative Smart Grid Technologies (ISGT)*. IEEE, 2012, pp. 1–8. [Online]. Available: http://dx.doi.org/10.1109/ISGT.2012.6175683

A. K. Chakraborty and N. Shaniia, "Advanced metering infrastructure: Technology and challenges," in *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T D)*, May 2016, pp. 1–5. [Online]. Available: http://dx.doi.org/10.1109/TDC.2016.7520076

S.-F. Chang, C.-F. Chen, J.-H. Wen, J.-H. Liu, J.-H. Weng, and J.-L. Dong, "Application and development of ZigBee technology for smart grid environment," *Journal of Power and Energy Engineering*, vol. 3, no. 04, p. 356, 2015. [Online]. Available: http://dx.doi.org/10.4236/jpee.2015.34048

D. Chen, J. Brown, and J. Y. Khan, "6LoWPAN based neighborhood area network for a smart grid communication infrastructure," in *2013 Fifth International Conference on Ubiquitous and Future Networks (ICUFN)*. Institute of Electrical & Electronics Engineers (IEEE), jul 2013. [Online]. Available: http://dx.doi.org/10.1109/icufn.2013.6614885

——, "Performance analysis of a distributed 6LoWPAN network for the smart grid applications," in *2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*. Institute of Electrical & Electronics Engineers (IEEE), apr 2014. [Online]. Available: http://dx.doi.org/10.1109/issnip.2014.6827646

X. Chen, Z. Dai, W. Li, and H. Shi, "A layer-based routing protocol for heterogeneous wireless sensor networks," in *2012 IEEE International Conference on Communications*

(ICC). Institute of Electrical & Electronics Engineers (IEEE), jun 2012. [Online]. Available: http://dx.doi.org/10.1109/icc.2012.6363670

B.-C. Cheng, G.-T. Liao, Y.-T. Liu, P.-H. Hsu, W. Tang, and Y.-H. Chen, "TIM-AODV: A secure routing protocol for preventing black hole attacks in real-time AMI smart meter networks," vol. 1, Orlando, FL, United states, 2013, pp. 268 – 273. [Online]. Available: http://www.iiis.org/CDs2013/CD2013SCI/SCI_2013/PapersPdf/SA971HW.pdf

H. R. Chi, K. F. Tsang, K. T. Chui, H. S.-H. Chung, B. W. K. Ling, and L. L. Lai, "Interference-mitigated ZigBee-based advanced metering infrastructure," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 672–684, 2016. [Online]. Available: http://dx.doi.org/10.1109/TII.2016.2527618

S. Chren, B. Rossi, and T. Pitner, "Smart grids deployments within EU projects: The role of smart meters," in *2016 Smart Cities Symposium Prague (SCSP)*, May 2016, pp. 1–5. [Online]. Available: http://dx.doi.org/10.1109/SCSP.2016.7501033

I. Colak, G. Fulli, S. Sagiroglu, M. Yesilbudak, and C.-F. Covrig, "Smart grid projects in europe: Current status, maturity and future scenarios," *Applied Energy*, vol. 152, pp. 58–70, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.apenergy.2015.04.098

H. Das and L. Saikia, "GSM enabled smart energy meter and automation of home appliances," in *2015 International Conference on Energy, Power and Environment: Towards Sustainable Growth (ICEPE)*. IEEE, 2015, pp. 1–5. [Online]. Available: http://dx.doi.org/10.1109/EPETSG.2015.7510071

M. El-Hawary, "The smart grid: State-of-the-art and future trends," *Electric Power Components and Systems*, vol. 42, no. 3-4, pp. 239 – 250, 2014. [Online]. Available: http://dx.doi.org/10.1080/15325008.2013.868558

S. Elyengui, R. Bouhouchi, and T. Ezzedine, "A comparative performance study of the routing protocols RPL, LOADng and LOADng-CTP with bidirectional traffic for AMI scenario," in *2015 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)*, Oct 2015, pp. 43–49. [Online]. Available: http://dx.doi.org/10.1109/ICSGCE.2015.7454267

M. Erol-Kantarci and H. T. Mouftah, "Energy-efficient information and communication infrastructures in the smart grid: A survey on interactions and open issues," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 179–197, 2015. [Online]. Available: http://dx.doi.org/10.1109/COMST.2014.2341600

H. Farhangi, "A road map to integration: Perspectives on smart grid development," *Power and Energy Magazine, IEEE*, vol. 12, no. 3, pp. 52–66, May 2014. [Online]. Available: http://dx.doi.org/10.1109/MPE.2014.2301515

H. Farooq and L. T. Jung, "Multi metric on-demand ad-hoc routing protocol for wireless smart metering deployment," Kuala Lumpur, Malaysia, 2013, pp. 140 – 144. [Online]. Available: http://dx.doi.org/10.1109/MICC.2013.6805814

——, "Performance analysis of ad-hoc routing protocols in smart metering infrastructure," London, United kingdom, 2013, pp. 871 – 874. [Online]. Available: http://ieeexplore.ieee.org/document/6661844/

H. A. Foudeh and A. S. Mokhtar, "Automated meter reading and advanced metering infrastructure projects," in *2015 9th Jordanian International Electrical and Electronics Engineering Conference (JIEEEC)*, Oct 2015, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1109/JIEEEC.2015.7470753

S. Galli, A. Scaglione, and Z. Wang, "For the grid and through the grid: The role of power line communications in the smart grid," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 998–1027, 2011. [Online]. Available: http://doi.acm.org/10.1109/jproc.2011.2109670

J. Garcia-Hernandez, "Recent progress in the implementation of AMI projects: Standards and communications technologies," in *Mechatronics, Electronics and Automotive Engineering (ICMEAE), 2015 International Conference on*, Nov 2015, pp. 251–256. [Online]. Available: http://dx.doi.org/10.1109/ICMEAE.2015.43

A. Ghasemkhani, A. Anvari-Moghaddam, J. M. Guerrero, and B. Bak-Jensen, "An efficient multi-objective approach for designing of communication interfaces in smart grids," in *IEEE PES, Innovative Smart Grid Technologies (ISGT 2016)*. IEEE Press, 2016. [Online]. Available: http://vbn.aau.dk/files/237972970/PID4381335.pdf

I. Gitman, "On the capacity of slotted ALOHA networks and some design problems," *Communications, IEEE Transactions on*, vol. 23, no. 3, pp. 305–317, 1975. [Online]. Available: http://dx.doi.org/10.1109/TCOM.1975.1092805

S. Gormus, Z. Fan, Z. Bocus, and P. Kulkarni, "Opportunistic communications to improve reliability of AMI mesh networks," in *Innovative Smart Grid Technologies (ISGT Europe), 2011 2nd IEEE PES International Conference and Exhibition on*. IEEE, 2011, Conference Proceedings, pp. 1–8. [Online]. Available: http://dx.doi.org/10.1109/ISGTEurope.2011.6162817

S. Gormus, F. Tosato, Z. Fan, Z. Bocus, and P. Kulkarni, "Opportunistic RPL for reliable AMI mesh networks," *Wireless Networks*, vol. 20, no. 8, pp. 2147–2164, may 2014. [Online]. Available: http://dx.doi.org/10.1007/s11276-014-0730-1

D. Grochocki, J. H. Huh, R. Berthier, R. Bobba, W. H. Sanders, A. A. Cárdenas, and J. G. Jetcheva, "AMI threats, intrusion detection requirements and deployment recommendations," in *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*, Nov 2012, pp. 395–400. [Online]. Available: http://dx.doi.org/10.1109/SmartGridComm.2012.6486016

V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "Smart grid technologies: communication technologies and standards," *Industrial informatics, IEEE transactions on*, vol. 7, no. 4, pp. 529–539, 2011. [Online]. Available: http://dx.doi.org/10.1109/TII.2011.2166794

V. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. Hancke, "A survey on smart grid potential applications and communication requirements," *Industrial Informatics, IEEE Transactions on*, vol. 9, no. 1, pp. 28–42, Feb 2013. [Online]. Available: http://dx.doi.org/10.1109/TII.2012.2218253

P. Gupta and P. R. Kumar, "The capacity of wireless networks," *Information Theory, IEEE Transactions on*, vol. 46, no. 2, pp. 388–404, 2000. [Online]. Available: http://dx.doi.org/10.1109/18.825799

T. Hartmann, F. Fouquet, J. Klein, Y. L. Traon, A. Pelov, L. Toutain, and T. Ropitault, "Generating realistic smart grid communication topologies based on real-data," in *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*, Nov 2014, pp. 428–433. [Online]. Available: http://dx.doi.org/10.1109/SmartGridComm.2014.7007684

R. Hassan and G. Radman, "Survey on smart grid," in *IEEE SoutheastCon 2010 (SoutheastCon), Proceedings of the*, March 2010, pp. 210–213. [Online]. Available: http://dx.doi.org/10.1109/SECON.2010.5453886

M. Hossain, N. Madlool, N. Rahim, J. Selvaraj, A. Pandey, and A. F. Khan, "Role of smart grid in renewable energy: An overview," *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 1168–1184, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.rser.2015.09.098

H. Hu, D. Kaleshi, A. Doufexi, and L. Li, "Performance analysis of ieee 802.11af standard based neighbourhood area network for smart grid applications," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5. [Online]. Available: http://dx.doi.org/10.1109/VTCSpring.2015.7146000

Hydro-Québec, "Remote meter-reading project phase 1," online, "Régie de l'énergie du Québec", May 2012, http://publicsde.regie-energie.qc.ca/projets/34/DocPrj/R-3770-2011-B-0157-TRAD-DOC-2012_05_23.pdf.

M. Islam, M. M. Uddin, M. A. A. Mamun, and M. Kader, "Performance analysis of AMI distributed area network using WiMAX technology," in *2014 9th International Forum on Strategic Technology (IFOST)*. Institute of Electrical & Electronics Engineers (IEEE), oct 2014. [Online]. Available: http://dx.doi.org/10.1109/ifost.2014.6991093

G. Iyer, P. Agrawal, E. Monnerie, and R. Cardozo, "Performance analysis of wireless mesh routing protocols for smart utility networks," in *Smart Grid Communications, 2011 IEEE International Conference on*, Oct 2011, pp. 114–119. [Online]. Available: http://dx.doi.org/10.1109/SmartGridComm.2011.6102301

G. Iyer, P. Agrawal, and R. S. Cardozo, "Performance comparison of routing protocols over smart utility networks: A simulation study," in *2013 IEEE Globecom Workshops*. Institute of Electrical & Electronics Engineers (IEEE), dec 2013. [Online]. Available: http://dx.doi.org/10.1109/glocomw.2013.6825116

K. Jain, J. Padhye, V. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," *Wireless networks*, vol. 11, no. 4, pp. 471–487, 2005. [Online]. Available: http://dx.doi.org/10.1007/s11276-005-1769-9

D. Jin, C. Lee, D. Nicol, I. Shin, and H. Zhu, "Simulation-based study of distributed denial-of-service attacks in advanced metering infrastructure," in *INFORMS Annual Meeting, Charlotte, NC, USA*, 2011.

D. Johnson, Y. Hu, and D. Maltz, "The dynamic source routing protocol (DSR) for mobile ad hoc networks for IPv4," *RFC4728*, pp. 2–100, 2007. [Online]. Available: http://dx.doi.org/10.17487/RFC4728

Y. Kabalci, "A survey on smart metering and smart grid communication," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 302–318, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.rser.2015.12.114

V. Kathuria, G. Mohanasundaram, and S. R. Das, "A simulation study of routing protocols for smart meter networks," Vancouver, BC, Canada, 2013, pp. 384 – 389. [Online]. Available: http://dx.doi.org/10.1109/SmartGridComm.2013.6687988

M. Kaynia and N. Jindal, "Performance of ALOHA and CSMA in spatially distributed wireless networks," in *Communications, 2008. ICC '08. IEEE International Conf. on*, May 2008, Journal Article, pp. 1108–1112. [Online]. Available: http://dx.doi.org/10.1109/icc.2008.216

A. A. Khan, M. H. Rehmani, and M. Reisslein, "Cognitive radio for smart grids: Survey of architectures, spectrum sensing mechanisms, and networking protocols," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 860–898, Firstquarter 2016. [Online]. Available: http://dx.doi.org/10.1109/COMST.2015.2481722

A. C. Kizilkale and R. P. Malhame, "Collective target tracking mean field control for markovian jump-driven models of electric water heating loads," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 1867–1872, 2014. [Online]. Available: http://dx.doi.org/10.3182/20140824-6-ZA-1003.00630

L. Kleinrock and S. Lam, "Packet switching in a multiaccess broadcast channel: Performance evaluation," *Communications, IEEE Transactions on*, vol. 23, no. 4, pp. 410–423, 1975. [Online]. Available: http://dx.doi.org/10.1109/TCOM.1975.1092814

M. Kodialam and T. Nandagopal, "Characterizing achievable rates in multi-hop wireless networks: the joint routing and scheduling problem," in *Proceedings of the 9th annual international conference on Mobile computing and networking*. ACM, 2003, Conference Proceedings, pp. 42–54. [Online]. Available: http://dx.doi.org/10.1145/938985.938991

P. Kulkarni, S. Gormus, and Z. Fan, "Tree balancing in smart grid advanced metering infrastructure mesh networks," in *Green Computing and Communications (GreenCom), 2012 IEEE International Conference on*. IEEE, 2012, Conference Proceedings, pp. 109–115. [Online]. Available: http://dx.doi.org/10.1109/GreenCom.2012.27

P. Kulkarni, S. Gormus, Z. Fan, and F. Ramos, "AMI mesh networks: a practical solution and its performance evaluation," *Smart Grid, IEEE Trans. on*, vol. 3, no. 3, pp. 1469–1481, 2012. [Online]. Available: http://dx.doi.org/10.1109/TSG.2012.2205023

N. T. Le and W. Benjapolakul, "Opportunistic hybrid network coding data gathering scheme for non-concurrent applications in smart grid neighborhood area network," in *2016 13th International Conference on Electrical Engineering/Electronics, Computer,*

Telecommunications and Information Technology (ECTI-CON), June 2016, pp. 1–5. [Online]. Available: http://dx.doi.org/10.1109/ECTICon.2016.7561249

S. Lee, J. Bong, S. Shin, and Y. Shin, "A security mechanism of smart grid AMI network through smart device mutual authentication," in *The International Conference on Information Networking 2014 (ICOIN2014)*, Feb 2014, pp. 592–595. [Online]. Available: http://dx.doi.org/10.1109/ICOIN.2014.6799750

X.-Y. Li, "Wireless transmission fundamentals," in *Wireless Ad Hoc and Sensor Networks*. Cambridge University Press, 2008, pp. 17–44, cambridge Books Online. [Online]. Available: http://dx.doi.org/10.1017/CBO9780511754722.003

B. Lichtensteiger, B. Bjelajac, C. Muller, and C. Wietfeld, "RF-mesh systems for smart metering: system architecture and performance," in *Smart Grid Communications, 2010 First IEEE Int. Conf. on*. IEEE, 2010, pp. 379–384. [Online]. Available: http://dx.doi.org/10.1109/SMARTGRID.2010.5622071

C. H. Lo and N. Ansari, "The progressive smart grid system from both power and communications aspects," *Communications Surveys & Tutorials, IEEE*, vol. 14, no. 3, pp. 799–821, 2012. [Online]. Available: http://dx.doi.org/10.1109/SURV.2011.072811.00089

G. Lopez, J. Moreno, H. Amaris, and F. Salazar, "Paving the road toward smart grids through large-scale advanced metering infrastructures," *Electric Power Systems Research*, vol. 120, pp. 194 – 205, 2015, smart Grids: World's Actual Implementations. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378779614001862

F. Malandra and B. Sansò, "Analytical performance analysis of a large-scale RF-mesh smart meter communication system," in *Innovative Smart Grid Technologies (ISGT North America), 2015 IEEE PES Conference on*, 2015, to appear. Available online at https://www.gerad.ca/en/papers/G-2015-46.

——, "A simulation framework for network performance evaluation of large-scale RF-Mesh AMIs," *Simulation Modelling Practice and Theory, submitted to.*, 2016.

F. Malandra and B. Sansó, "PeRF-mesh: A performance analysis tool for large scale RF-mesh-based smart meter networks with FHSS," in *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm): Communications and Networks to Enable the Smart Grid (IEEE SmartGridComm'15 Symposium - Communications and Networks)*, Miami, USA, Nov. 2015. [Online]. Available: http://dx.doi.org/10.1109/SmartGridComm.2015.7436398

K. Masood, M. S. Sohail, and A. U. Sheikh, "Slotted ALOHA performance for FU-FB in frequency selective fading environment," *Wireless Communications and Mobile Computing*, vol. 15, no. 2, pp. 228–242, 2015. [Online]. Available: http://dx.doi.org/10.1002/wcm.2340

M. Médard and A. J. Goldsmith, "Capacity of time-slotted ALOHA systems," in *Wireless Communications and Networking Conference, 1999. WCNC. 1999 IEEE.* IEEE, Conference Proc., pp. 490–494.

A. Michaloliakos, R. Rogalin, Y. Zhang, K. Psounis, and G. Caire, "Performance modeling of next-generation WiFi networks," *Computer Networks*, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2016.05.024

R. R. Mohassel, A. S. Fung, F. Mohammadi, and K. Raahemifar, "A survey on advanced metering infrastructure and its application in smart grids," in *Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on*, May 2014, pp. 1–8. [Online]. Available: http://dx.doi.org/10.1109/CCECE.2014.6901102

Municipality of Montréal, "List of residential addresses," 2014, Portail données ouvertes Montréal, data available at http://donnees.ville.montreal.qc.ca/dataset/adresses-ponctuelles.

V. Namboodiri, V. Aravinthan, and W. Jewell, "Communication needs and integration options for AMI in the smart grid," Power Systems Engineering Research Center, Tech. Rep., 03 2012. [Online]. Available: http://www.pserc.wisc.edu/documents/publications/papers/fgwhitepapers/Namboodiri_PSERC_Future_Grid_White_Paper_AMI_May_2012.pdf

T. Nandagopal, T. E. Kim, X. Gao, and V. Bharghavan, "Achieving MAC layer fairness in wireless packet networks," in *Proceedings of the 6th annual international conference on Mobile computing and networking.* ACM, Conference Proceedings, pp. 87–98. [Online]. Available: http://dx.doi.org/10.1145/345910.345925

S. Narvare and R. D. M. Shrivastava, "Review paper of performance analysis for random access channel in wireless network with MAC protocol," *Int. Journal of Advanced Tech. and Engineering Exploration (IJATEE)*, vol. 1, pp. 7–14, 2014.

Navigant Research, "Global AMI tracker 2Q16. Smart Meter Projects: Project Tracking, Regional Analysis, and Vendor Selection Share," Tech. Rep., 06 2016. [Online]. Available: http://www.navigantresearch.com/research/global-ami-tracker-2q16

NIST, "NIST framework and roadmap for smart grid interoperability standards," NIST, Tech. Rep. NIST Special Publication 1108R2, 2010. [Online]. Available: https://www.nist.gov/document-7269

A. Noorwali, R. Rao, and A. Shami, "Modelling and delay analysis of wireless home area networks in a smart grid," in *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Nov 2015, pp. 569–574. [Online]. Available: http://dx.doi.org/10.1109/SmartGridComm.2015.7436361

M. K. Oh, S. Choi, and C. h. Shin, "Narrowband PHY design and its transmission scheme for AMI applications," in *Information and Communication Technology Convergence (ICTC), 2015 International Conference on*, Oct 2015, pp. 1197–1199. [Online]. Available: http://dx.doi.org/10.1109/ICTC.2015.7354774

S. Panchadcharam, G. A. Taylor, Q. Ni, I. Pisica, and S. Fateri, "Performance evaluation of smart metering infrastructure using simulation tool," in *2012 47th International Universities Power Engineering Conference (UPEC)*. Institute of Electrical & Electronics Engineers (IEEE), sep 2012. [Online]. Available: http://dx.doi.org/10.1109/upec.2012.6398571

I. Parvez, N. Islam, N. Rupasinghe, A. I. Sarwat, and I. Guvenc, "LAA-based LTE and ZigBee coexistence for unlicensed-band smart grid communications," in *SoutheastCon 2016*, March 2016, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1109/SECON.2016.7506714

A. Patel, J. Aparicio, N. Tas, M. Loiacono, and J. Rosca, "Assessing communications technology options for smart grid applications," in *Smart Grid Communications, 2011 IEEE International Conference on*. IEEE, 2011, pp. 126–131. [Online]. Available: http://dx.doi.org/10.1109/SmartGridComm.2011.6102303

C. Peng and J. Huang, "A home energy monitoring and control system based on ZigBee technology," *International Journal of Green Energy*, no. just-accepted, 2016. [Online]. Available: http://dx.doi.org/10.1080/15435075.2016.1188102

C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *Mobile Computing Systems and Applications, 1999. Proceedings. WMCSA'99. Second IEEE Workshop on*. IEEE, 1999, Conference Proceedings, pp. 90–100. [Online]. Available: http://dx.doi.org/10.17487/RFC3561

H. J. Pozveh, H. Mohammadinejad, and M. Bateni, "Improving the AODV protocol to satisfy the required level of reliability for home area networks." *International Journal*

of Computer Network & Information Security, vol. 8, no. 6, 2016. [Online]. Available: http://dx.doi.org/10.5815/ijcnis.2016.06.03

D. F. Ramirez, S. Cespedes, C. Becerra, and C. Lazo, "Performance evaluation of future AMI applications in smart grid neighborhood area networks," in *Communications and Computing (COLCOM), 2015 IEEE Colombian Conference on*. IEEE, 2015, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1109/ColComCon.2015.7152093

J. R. R. Renofio, M. E. Pellenz, A. Santin, E. Jamhour, M. C. Penna, and R. D. Souza, "Insights on the resilience and capacity of AMI wireless networks," in *2016 IEEE Symposium on Computers and Communication (ISCC)*, June 2016, pp. 610–615. [Online]. Available: http://dx.doi.org/10.1109/ISCC.2016.7543804

D. T. Riascos, E. I. Ortega, and A. A. Tellez, "Advanced metering infrastructure for microgeneration in rural zones using tv white spaces based on spectrum opportunistic," *IEEE Latin America Transactions*, vol. 13, no. 12, pp. 3813–3819, Dec 2015. [Online]. Available: http://dx.doi.org/10.1109/TLA.2015.7404913

E. B. Rice and A. AlMajali, "Mitigating the risk of cyber attack on smart grid systems," *Procedia Computer Science*, vol. 28, pp. 575–582, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.procs.2014.03.070

C. C. Robusto, "The cosine-haversine formula," *The American Mathematical Monthly*, vol. 64, no. 1, p. 38, jan 1957. [Online]. Available: http://dx.doi.org/10.2307/2309088

R. Rom and M. Sidi, *Multiple access protocols: performance and analysis*. Springer Science & Business Media, 2012.

A. I. Sabbah, A. El-Mougy, and M. Ibnkahla, "A survey of networking challenges and routing protocols in smart grids," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 210–221, Feb 2014. [Online]. Available: http://dx.doi.org/10.1109/TII.2013.2258930

N. Saputro, K. Akkaya, and S. Uludag, "A survey of routing protocols for smart grid communications," *Computer Networks*, vol. 56, no. 11, pp. 2742–2771, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2012.03.027

S. Shao, S. Guo, X. Qiu, L. Meng, and M. Lei, "Traffic scheduling mechanism based on interference avoidance for meter data collection in wireless smart grid communication networks," *China Communications*, vol. 12, no. 7, pp. 142–153, July 2015. [Online]. Available: http://dx.doi.org/10.1109/CC.2015.7188532

K. Sharma and L. M. Saini, “Performance analysis of smart metering for smart grid: An overview,” *Renewable and Sustainable Energy Reviews*, vol. 49, pp. 720–735, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.rser.2015.04.170

J. A. Silvester and L. Kleinrock, “On the capacity of multihop slotted ALOHA networks with regular structure,” *Communications, IEEE Transactions on*, vol. 31, no. 8, pp. 974–982, 1983. [Online]. Available: http://dx.doi.org/10.1109/TCOM.1983.1095921

F. Sirois, B. Bourdel, and R. Malhamé, “Projet rene-034: Gestion des capacités de stockage d’énergie dispersées dans un réseau électrique pour l’amortissement des effets de variabilité des sources d’énergie renouvelables,” Public report, 2016.

J. Solis, “Développement d’un estimateur d’état énergétique d’un chaffe-eau pour un contrôle par champ mpyen,” Master’s thesis, 2015. [Online]. Available: https://publications.polymtl.ca/2014/

W. Somkaew, S. Thepphaeng, and C. Pirak, “Data security implementation over ZigBee networks for AMI systems,” in *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2014 11th International Conference on*, May 2014, pp. 1–5. [Online]. Available: http://dx.doi.org/10.1109/ECTICon.2014.6839872

C. Sousa, G. e Souza, I. Moraes, R. C. Carrano, C. V. N. de Albuquerque, L. F. N. Passos, A. Carniato, A. L. Bettiol, R. Z. Homma, R. C. Andrade, and F. H. Molina, “Link quality estimation for AMI,” in *Innovative Smart Grid Technologies Latin America (ISGT LATAM), 2015 IEEE PES*, Oct 2015, pp. 646–649. [Online]. Available: http://dx.doi.org/10.1109/ISGT-LA.2015.7381231

E. U. Soykan, S. D. Ersoz, and G. Soykan, “Identity based signcryption for advanced metering infrastructure,” in *Smart Grid Congress and Fair (ICSG), 2015 3rd International Istanbul*, April 2015, pp. 1–5. [Online]. Available: http://dx.doi.org/10.1109/SGCF.2015.7354933

A. Tammam, “Lissage optimal de la charge électrique en présence de sources d’énergies renouvelables via le pilotage de la consommation des chauffe-eau,” PhD dissertation, Ecole Polytechnique de Montréal, 2016.

M. S. Thomas, I. Ali, and N. Gupta, “A secure way of exchanging the secret keys in advanced metering infrastructure,” in *Power System Technology (POWERCON), 2012 IEEE International Conference on*, Oct 2012, pp. 1–7. [Online]. Available: http://dx.doi.org/10.1109/PowerCon.2012.6401418

S. Tonyali, K. Akkaya, N. Saputro, and A. S. Uluagac, "A reliable data aggregation mechanism with homomorphic encryption in smart grid AMI networks," in *2016 13th IEEE Annual Consumer Communications Networking Conference (CCNC)*, Jan 2016, pp. 550–555. [Online]. Available: http://dx.doi.org/10.1109/CCNC.2016.7444839

J. Tripathi, J. C. de Oliveira, and J. P. Vasseur, "A performance evaluation study of RPL: routing protocol for low power and lossy networks," in *Information Sciences and Systems (CISS), 2010 44th Annual Conference on*. IEEE, 2010, Conference Proceedings, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1109/CISS.2010.5464820

M. L. Tuballa and M. L. Abundo, "A review of the development of smart grid technologies," *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 710–725, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.rser.2016.01.011

US DoE, "Communication requirements of smart grid technologies," U.S. Department of Energy), Report, 2010. [Online]. Available: http://energy.gov/sites/prod/files/gcprod/documents/Smart_Grid_Communications_Requirements_Report_10-05-2010.pdf

——, "Implementing the national broadband plan by studying the communications requirements of electric utilities to inform federal smart grid policy," *Federal Register*, vol. 75, no. 90, p. 26206, May 2010, available at https://federalregister.gov/a/2010-11129.

L. Vangelista, A. Zanella, and M. Zorzi, "Long-range IoT technologies: The dawn of LoRa," in *Future Access Enablers of Ubiquitous and Intelligent Infrastructures*. Springer, 2015, pp. 51–58. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-27072-2_7

D. Wang, Z. Tao, J. Zhang, and A. A. Abouzeid, "RPL based routing for advanced metering infrastructure in smart grid," in *Communications Workshops (ICC), 2010 IEEE International Conference on*. IEEE, 2010, Conference Proceedings, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1109/ICCW.2010.5503924

H. Wang, M. Xi, J. Liu, and C. Chen, "Transmitting IPv6 packets over bluetooth low energy based on BlueZ," in *Advanced Communication Technology (ICACT), 2013 15th International Conference on*. IEEE, 2013, pp. 72–77. [Online]. Available: http://ieeexplore.ieee.org/document/6488142/

M. Xiang, Q. Bai, and W. Liu, "Self-adjustable trust-based energy efficient routing for smart grid systems," in *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 03*.

IEEE Computer Society, 2012, Conference Proceedings, pp. 378–382. [Online]. Available: http://dx.doi.org/10.1109/WI-IAT.2012.89

E. Yaacoub and A. Kadri, "LTE radio resource management for real-time smart meter reading in the smart grid," in *2015 IEEE International Conference on Communication Workshop (ICCW)*. IEEE, 2015, pp. 2000–2005. [Online]. Available: http://dx.doi.org/10.1109/ICCW.2015.7247474

M. Yigit, E. A. Yoney, and V. C. Gungor, "Performance of MAC protocols for wireless sensor networks in harsh smart grid environment," in *2013 First International Black Sea Conference on Communications and Networking (BlackSeaCom)*. Institute of Electrical & Electronics Engineers (IEEE), jul 2013. [Online]. Available: http://dx.doi.org/10.1109/blackseacom.2013.6623380

T. Zahariadis, H. Leligou, S. Voliotis, S. Maniatis, P. Trakadas, and P. Karkazis, "An energy and trust-aware routing protocol for large wireless sensor networks," in *WSEAS international conference on Applied informatics and communications (AIC'09). World Scientific and Engineering Academy and Society (WSEAS)*, 2009. [Online]. Available: http://www.ee.teihal.gr/labs/digilab_site/AIC38.pdf

M. Zareei, A. Zarei, R. Budiarto, and M. Omar, "A comparative study of short range wireless sensor network on high density networks," in *Communications (APCC), 2011 17th Asia-Pacific Conference on*, Oct 2011, pp. 247–252. [Online]. Available: http://dx.doi.org/10.1109/APCC.2011.6152813

Q. Zeng, H. Li, and D. Peng, "Frequency-hopping based communication network with multi-level qoss in smart grid: Code design and performance analysis," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1841–1852, Dec 2012. [Online]. Available: http://dx.doi.org/10.1109/TSG.2012.2214067

K. Zheng, F. Liu, L. Lei, C. Lin, and Y. Jiang, "Stochastic performance analysis of a wireless finite-state markov channel," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 782–793, February 2013. [Online]. Available: http://dx.doi.org/10.1109/TWC.2012.122212.120223

APPENDIX A Poisson-distributed traffic characterization

Throughout this work the assumption of a Poisson-distributed traffic characterization is widely adopted. The main advantage of using this type of traffic distribution is the possibility to aggregate multiple traffic streams. One of the basic properties of Poisson-distributed random variables is that the aggregation of two or more Poisson-distributed variables with mean parameter λ_i is still Poisson-distributed with mean parameter $\sum_i \lambda_i$.

This has a great impact in the mathematical modelling because it provides an easy way to define the traffic generation rate in multi-hop networks with a static routing protocol and multiple traffic sources. The traffic characterization of the three types of devices in the architecture analyzed in this thesis is illustrated in Figure 5.1 in Chapter 5.

The routing mechanism is based on the shortest path between each smart meter and its associated data collector. Let ξ_i be the number of shortest paths that includes node i as intermediate node.

In Figure A.1, the traffic distribution of smart meters is illustrated: smart meters can be sources, destinations or intermediate nodes in a shortest path. Black arrows indicate the uplink direction, whereas blue arrows the downlink direction. In the uplink, a packet directed to the data collector is generated by each smart meter according to a Poisson distribution of mean parameter λ_{up}. As one can see, smart meter i receives at a rate $\xi_i\lambda_{up}$ from the ξ_i smart meters for which it is an intermediate node. On top of that, it receives at a rate λ_{up} from the HAN. Therefore, it transmits at a rate $(\xi_i + 1)\lambda_{up}$ in the uplink direction. On the other hand, in the downlink direction, a data collector produces packets directed to each

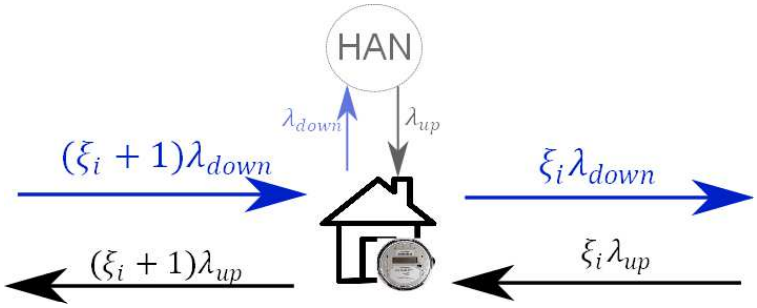


Figure A.1 Poisson-distributed traffic characterization of smart meters under the hypotheses of static routing

smart meter at a rate λ_{down}. Smart meter i receives at a rate $(\xi_i + 1)\lambda_{down}$: of which, λ_{down} is directed to its HAN and $\xi_i\lambda_{down}$ is forwarded along the ξ_i shortest paths.

In Figure A.2, the traffic distribution of routers is illustrated: routers can only be intermediate nodes in a shortest path, since they do not generate any packets and no packet is directed to them. They receive and transmit at a rate of $\xi_i\lambda_{down}$ in the downlink direction (blue arrows), and at a rate of $\xi_i\lambda_{up}$ in the uplink direction (black arrows).

In Figure A.3, the traffic distribution of data collectors is illustrated: data collectors can only be sources or destinations in a shortest path. They generate packets directed to each smart meter at a rate λ_{down}, leading to an aggregate rate of $|M|\lambda_{down}$ being $|M|$ the number of smart meters. On the other hand, they receive an aggregate traffic of $|M|\lambda_{up}$.

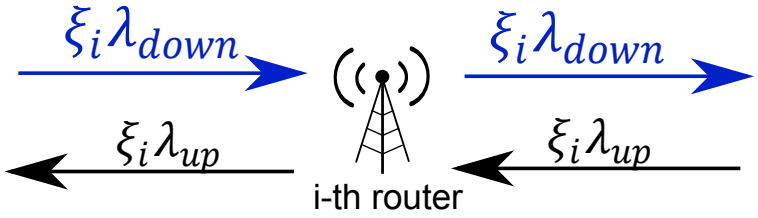


Figure A.2 Poisson-distributed traffic characterization of routers under the hypotheses of static routing

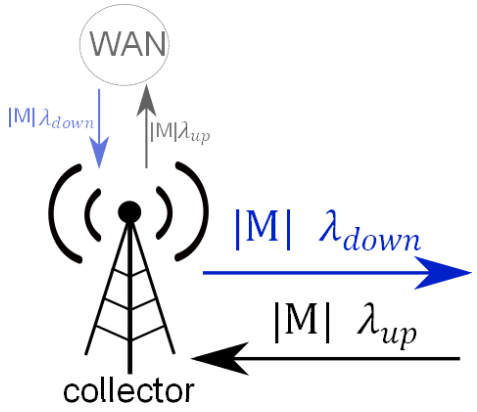


Figure A.3 Poisson-distributed traffic characterization of data collectors under the hypotheses of static routing

APPENDIX B Additional mathematical proofs for the Markov-modulated model in Chapter 7

In the Markov-modulated model presented in Chapter 7, 3 sequential phases (i.e. *Beginning*, *Transmission*, and *End*) are adopted in order to describe the operating state of a node in each time slot.

The sets of all the possible states of node $i \in V$ in the *Beginning*, *Transmission*, and *End* phase of time slot t are denoted with $S_B(i,t)$, $S_T(i,t)$, and $S_E(i,t)$ respectively.

The state probabilities for all the nodes $i \in V$ are denoted as follows:

- $\pi_B(i,n,t)$ is the probability that node i is in the *Beginning* phase with n packets in the buffer at time t
- $\pi_T(i,\varnothing,n,t)$ is the probability that node i is idle in the *Transmission* phase with n packets in the buffer at time t
- $\pi_T(i,j,n,t)$ is the probability that node i is transmitting to node j with n packets in the buffer at time t
- $\pi_E(i,n,t)$ is the probability that node i is in the beginning phase with n packets in the buffer at time t

Let Z be the buffer size, $\Gamma(i)$ be the set of neighbors of node i, and $\alpha_i(\Gamma_k)$ be the percentage of traffic that node i transmits to its k-th neighbor Γ_k. Since $\alpha_i(\Gamma_k)$ is defined as a percentage, it implies that:

$$\sum_{j \in \Gamma(i)} \alpha_i(j) = 1 \tag{B.1}$$

The beginning state probabilities at time $t = 1$ are chosen as initial values, then the model is iteratively solved, as indicated in Section 7.4. The integrity of the model is based on the following theorem:

Theorem 1. *Each node sequentially undergoes the three defined phases (i.e., Beginning, Transmission, and End) at each time slot.*

In order to prove Theorem 1, we define and prove the three following lemmas.

Lemma 2. *If a node is in the Beginning phase at time t it will evolve to the Transmission phase in the same time slot with probability equal to one:*

$$\sum_{S\in S_B(i,t)} \pi_S = 1 \implies \sum_{S\in S_T(i,t)} \pi_S = 1 \quad \forall i \in V, \forall t.$$

Proof.

$$\begin{aligned}
\sum_{S\in S_B(i,t)} \pi_S &= 1 && \text{Initial assumption.}\\
\pi_B(i,0,t) + \sum_{n=1}^{Z} \pi_B(i,n,t) &= 1 &&\\
\pi_T(i,\varnothing,0,t) + \sum_{n=1}^{Z} (\pi_B(i,n,t)p_r + \pi_B(i,n,t)(1-p_r)) &= 1 && \text{See Equation (7.6).}\\
\pi_T(i,\varnothing,0,t) + \sum_{n=1}^{Z} p_r\pi_B(i,n,t) + \sum_{n=1}^{Z} \pi_T(i,\varnothing,n,t) &= 1 && \text{See Equation (7.8).}\\
\pi_T(i,\varnothing,0,t) + \sum_{n=1}^{Z} \left(p_r\pi_B(i,n,t) \sum_{j\in\Gamma(i)} \alpha_i(j) \right) + \sum_{n=1}^{Z} \pi_T(i,\varnothing,n,t) &= 1 && \text{See Equation (B.1).}\\
\pi_T(i,\varnothing,0,t) + \sum_{n=1}^{Z} \sum_{j\in\Gamma(i)} (p_r\pi_B(i,n,t)\alpha_i(j)) + \sum_{n=1}^{Z} \pi_T(i,\varnothing,n,t) &= 1 &&\\
\pi_T(i,\varnothing,0,t) + \sum_{n=1}^{Z} \sum_{j\in\Gamma(i)} (\pi_T(i,j,n,t))) + \sum_{n=1}^{Z} \pi_T(i,\varnothing,n,t) &= 1 && \text{See Equation (7.7).}\\
\pi_T(i,\varnothing,0,t) + \sum_{n=1}^{Z} \left(\pi_T(i,\varnothing,n,t) + \sum_{j\in\Gamma(i)} \pi_T(i,j,n,t)) \right) &= 1 &&\\
\sum_{S\in S_T(i,t)} \pi_S &= 1 && \square
\end{aligned}$$

Lemma 3. *If a node is in the Transmission phase at time t it will evolve to the End phase in the same time slot with probability equal to one:*

$$\sum_{S\in S_T(i,t)} \pi_S = 1 \implies \sum_{S\in S_E(i,t)} \pi_S = 1 \quad \forall i \in V, \forall t$$

Proof. According to the initial assumption of Lemma 3, we have:

$$\sum_{S\in S_T(i,t)} \pi_S = \sum_{n=0}^{Z} \pi_T(i,\varnothing,n,t) + \sum_{n=1}^{Z} \left(\sum_{\Gamma_k\in\Gamma(i)} \pi_T(i,\Gamma_k,n,t)) \right) = 1 \quad \forall i \in V, \forall t \qquad \text{(B.2)}$$

Moreover, if we sum up all the terms of Equations (7.15), (7.16), and (7.17) we obtain:

$$\begin{aligned}
\sum_{n=0}^{Z} \pi_E(i,n,t) &= \sum_{n=0}^{Z} \pi_T(i,\varnothing,n,t) + \sum_{n=1}^{Z} \sum_{\Gamma_k \in \Gamma(i)} \pi_T(i,\Gamma_k,1,t)(1-\beta_i^k(t)) + \\
&\quad + \sum_{n=1}^{Z} \sum_{\Gamma_k \in \Gamma(i)} \pi_T(i,\Gamma_k,1,t)(\beta_i^k(t)) \\
&= \sum_{n=0}^{Z} \pi_T(i,\varnothing,n,t) + \sum_{n=1}^{Z} \left((1-\beta_i^k(t)+\beta_i^k(t)) \sum_{\Gamma_k \in \Gamma(i)} \pi_T(i,\Gamma_k,1,t) \right) \\
&= \sum_{n=0}^{Z} \pi_T(i,\varnothing,n,t) + \sum_{n=1}^{Z} \left(\sum_{\Gamma_k \in \Gamma(i)} \pi_T(i,\Gamma_k,1,t) \right) \\
&= 1
\end{aligned}$$

□

Lemma 4. *If a node is in the End phase at time t it will evolve to the Beginning phase at time $t+1$ with probability equal to one:*

$$\sum_{S \in S_E(i,t)} \pi_S = 1 \implies \sum_{S \in S_B(i,t+1)} \pi_S = 1 \quad \forall i \in V, \forall t$$

Proof. Let $S_E(i,t)$ and $S_B(i,t+1)$ be the sets of all the possible states in respectively the *End* phase of time t and the *Transmission phase* of time $t+1$. According to the initial assumption of Lemma 4, we have:

$$\sum_{S \in S_E(i,t)} \pi_S = \sum_{n=0}^{Z} \pi_E(i,n,t) = 1 \quad \forall i \in V, \forall t \tag{B.3}$$

Moreover, if we sum up all the terms of Equations (7.20), (7.20), and (7.20) we obtain:

$$\begin{aligned}\sum_{n=0}^{Z}\pi_B(i,n,t+1) &= (1-\gamma_i)\pi_E(i,0,t) + (1-\gamma_i)\sum_{n=1}^{Z-1}\pi_E(i,n,t) + \gamma_i\sum_{n=1}^{Z-1}\pi_E(i,n-1,t) + \\ &\quad + \pi_E(i,Z,t) + \gamma_i\pi_E(i,Z-1,t) \\ &= (1-\gamma_i)\pi_E(i,0,t) + (1-\gamma_i)\sum_{n=1}^{Z-1}\pi_E(i,n,t) + \gamma_i\sum_{n=0}^{Z-2}\pi_E(i,n,t) + \\ &\quad + \pi_E(i,Z,t) + \gamma_i\pi_E(i,Z-1,t) \\ &= \sum_{n=0}^{Z-1}\pi_E(i,n,t)(1-\gamma_i+\gamma_i) + \pi_E(i,Z,t) \\ &= \sum_{n=0}^{Z}\pi_E(i,n,t) \\ &= 1 \end{aligned}$$ □

The combination of Lemmas 2, 3, and 4 permits to prove that:

$$\sum_{S\in S_B(i,t)}\pi_S = 1 \implies \sum_{S\in S_T(i,t)}\pi_S = 1 \implies \sum_{S\in S_E(i,t)} \implies \sum_{S\in S_B(i,t+1)}\pi_S = 1 \quad \forall i\in V, \forall t \tag{B.4}$$

The model is iteratively solved starting from the beginning state probabilities at time slot $t = 1$ which are chosen as follows:

$$\begin{aligned}&\pi_B(i,0,1) = 1 \ \ \forall i \in V \\ &\pi_B(i,n,1) = 0 \ \ \forall n \in (0,Z], \forall i \in V\end{aligned} \tag{B.5}$$

One can notice that in the initial solution ($t = 1$) we have $\sum_{S\in S_B(i,1)}\pi_S = 1$ meaning that each node $i \in V$ is in the *Beginning* phase in the first time slot with probability equal to one. We can use Lemma 2 to prove that the node will pass to the *Transmission* phase with probability equal to one in the same time slot. Similarly, using Lemma 3, we can state that each node will evolve to the *End* phase in the same time slot with probability equal to one. Finally, we can use Lemma 4 to prove that each node will be in the *Beginning* phase in the following time slot with probability equal to one. Recursively applying this procedure, we can easily see that the state of the system will sequentially undergo the three defined phases, as stated in Theorem 1.

Theorem 1 is then verified for all the initial solutions for which $\sum_{S \in S_B(i,1)} \pi_S = 1$.